

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 07:28:32 ; Search time 203.368 Seconds  
(without alignments)  
116.704 Million cell updates/sec

Title: US-09-498-556C-59

Perfect score: 486

Sequence: 1 KATMCGENKEYDSCGSEK.....VSAEDCLNDMDFTVPGTRN 84

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*

- 1: Geneseqp1980s.\*
- 2: Geneseqp1990s.\*
- 3: Geneseqp2000s.\*
- 4: Geneseqp2001s.\*
- 5: Geneseqp2002s.\*
- 6: Geneseqp2003as.\*
- 7: Geneseqp2003bs.\*
- 8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	486	100.0	84	AAV30432	Mature ne
2	486	100.0	84	AAI15317	A. caninu
3	486	100.0	91	AAI15317	A. caninu
4	486	100.0	91	AAI15317	A. caninu
5	486	100.0	91	AAI15317	A. caninu
6	486	100.0	91	AAI15317	A. caninu
7	212	43.6	82	AAV30422	Mature ne
8	212	43.6	82	AAV30399	Nematode
9	212	43.6	82	AAI15293	A. ceylan
10	212	43.6	82	AAI15307	A. ceylan
11	212	43.6	171	AAV30435	Mature ne
12	212	43.6	190	AAI15317	A. caninu
13	212	43.6	190	AAV30384	Nematode
14	202	41.6	86	AAV30419	Mature ne
15	202	41.6	86	AAI15304	A. caninu
16	202	41.6	105	AAI15308	A. caninu
17	202	41.6	105	AAV30408	Nematode
18	201	41.4	89	AAV30421	Mature ne
19	201	41.4	108	AAV30398	Nematode
20	201	41.4	108	AAI15292	A. ceylan
21	199	40.9	86	AAV30418	Mature ne
22	199	40.9	86	AAI15303	A. caninu
23	199	40.9	105	AAI151707	AcanAP25.
24	199	40.9	105	AAV30406	Nematode
25	197.5	40.6	88	AAV30416	Mature ne

26	197.5	40.6	88	3	AAI15301	A. caninu
27	197.5	40.6	107	2	AAI15301	A. caninu
28	197.5	40.6	107	2	AAI15301	A. caninu
29	195.5	40.2	84	2	AAV30423	Mature ne
30	195.5	40.2	162	3	AAV30436	Mature ne
31	195.5	40.2	162	3	AAI15321	A. caninu
32	195.5	40.2	181	2	AAI15321	A. caninu
33	195.5	40.2	181	2	AAI15321	A. caninu
34	195.5	40.2	181	2	AAI15321	A. caninu
35	193.5	39.8	162	3	AAI15322	A. caninu
36	193.5	39.8	162	3	AAI15322	A. caninu
37	193.5	39.8	181	2	AAI15322	A. caninu
38	193.5	39.8	181	2	AAI15322	A. caninu
39	186	38.3	83	2	AAV30425	Mature ne
40	186	38.3	102	2	AAV30401	Nematode
41	186	38.3	102	3	AAI15295	A. duoden
42	186	38.3	161	2	AAV30438	Mature ne
43	186	38.3	180	2	AAI15295	A. duoden
44	186	38.3	180	2	AAI15295	A. duoden
45	184.5	38.0	87	2	AAV30417	Mature ne

ALIGNMENTS

RESULT 1  
AAV30432  
ID AAV30432 standard; protein; 84 AA.  
AC AAV30432;  
XX  
DT 15-NOV-1999 (first entry)  
XX  
DE Mature nematode extracted anticoagulant protein AcanAPc2.  
XX  
KW Nematode extracted anticoagulant protein; NAP; anticoagulant;  
KW serine protease inhibitor; NAP domain; factor VIIa/TF.  
XX  
OS Ancylostoma caninum.  
XX  
PN US955294-A.  
XX  
PD 21-SEP-1999.  
XX  
PF 19-APR-1996; 96US-00634641.  
XX  
PR 18-OCT-1994; 94US-00326110.  
PR 05-JUN-1995; 95US-00461965.  
PR 05-JUN-1995; 95US-00465380.  
PR 05-JUN-1995; 95US-00486397.  
PR 05-JUN-1995; 95US-00486399.  
PR 17-OCT-1995; 95WO-US013231.  
XX  
(CORV-) CORVAS INT INC.  
XX  
PI Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;  
PI Bergum FW, Messens JHL, Laroche YR, Vlaeuk GP;  
XX  
WPI; 1999-539569/45.  
XX  
Screening an isolated protein for Nematode-extracted Anticoagulant  
Protein domains.  
XX  
PS Disclosure; Col 142-144; 197pp; English.  
XX  
The present sequence represents a nematode extracted anticoagulant  
protein (NAP). The protein has activity as an anticoagulant and/or serine  
protease inhibitor. The protein contains at least one NAP domain which  
has selective inhibitory activity for factor VIIa/TF. The specification  
describes a method for screening an isolated protein at least one domain  
for factor VIIa/TF selective inhibitory activity. The method comprises  
determining the time to clotting effected by a concentration of the  
isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo

CC activated partial thromboplastin time (APTT) assay; calculating  
 CC prolongation of clotting effected by the isolated protein in each of the  
 CC PT and APTT assay, with respect to a baseline clotting value for each  
 CC assay, where prolongation of clotting is calculated as fold elevation of  
 CC clotting time relative to a baseline clotting value, where a doubling of  
 CC clotting time is deemed a two-fold elevation; and calculating a PT to  
 CC APTT prolongation ratio, where a ratio at least one is indicative of  
 CC factor VIIa/TF inhibitory activity. The method is useful for determining  
 CC if a protein has factor VIIa/TF inhibitory activity  
 XX  
 SQ Sequence 84 AA;

Query Match 100.0%; Score 486; DB 2; Length 84;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-38;  
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KATWQCGENKYSKSGKCDKCKYDGVVEEDDEEPNVPCLVRVCHQDCVCEGFYRNK 60  
 DB 1 KATWQCGENKYSKSGKCDKCKYDGVVEEDDEEPNVPCLVRVCHQDCVCEGFYRNK 60  
 QY 61 DDKVSAEDCELDNMDFIYPGTRN 84  
 DB 61 DDKVSAEDCELDNMDFIYPGTRN 84

RESULT 2  
 AAB15317  
 ID AAB15317 standard; protein; 84 AA.  
 XX  
 AC AAB15317;  
 DT 06-AUG-2003 (revised)  
 DT 19-DEC-2000 (first entry)  
 XX  
 DE A. caninum nematode-extracted anticoagulant protein AcanAPc2(mature).  
 XX  
 KW Nematode-extracted anticoagulant protein; AcanAPc2; blood clotting;  
 KW canine hookworm; thrombosis; vaccine.  
 XX  
 OS Ancylostoma caninum.

XX US6087487-A.  
 XX 11-JUL-2000.  
 XX 12-FEB-1999; 99US-00249451.  
 XX 18-OCT-1994; 94US-00326110.  
 XX 05-JUN-1995; 95US-00461965.  
 XX 05-JUN-1995; 95US-00465380.  
 XX 05-JUN-1995; 95US-00486397.  
 XX 05-JUN-1995; 95US-00486399.  
 XX 17-OCT-1995; 95US-00486397.  
 XX 17-OCT-1995; 95US-00486399.  
 XX 17-APR-1997; 97US-00809455.  
 XX (CORV-) CORVAS INT INC.

XX Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;  
 XX Bergum PW, Messens JHL, Laroche YR, Vlasuk GP;  
 XX WPI; 2000-531359/48.  
 XX N-PSDB; AAA73373.  
 XX New cDNA molecule encoding a protein having factor Xa inhibitory activity  
 XX for preventing and treating blood clotting disorders, comprises nematode-  
 XX extracted anticoagulant protein domains.

XX Example A; Fig 16; 197pp; English.  
 XX The present sequence is the Ancylostoma caninum nematode-extracted  
 XX anticoagulant protein AcanAPc2. Proteins of this kind have been shown to  
 XX be effective at preventing blood clotting without causing excessive  
 XX bleeding. The protein can be used in blood collection tubes to aid the

CC isolation of plasma from the blood, to prevent thrombosis which may be  
 CC linked to the rupture of an atherosclerotic plaque, acute myocardial  
 CC infarction, angina, thrombolytic therapy, percutaneous transluminal  
 CC coronary angioplasty, disseminated intravascular coagulopathy, infection,  
 CC cancer and septic shock, and to produce antibodies. In the latter  
 CC instances, the antibodies can be raised in order to detect infection by  
 CC nematodes (the coding sequence can also be used for this) or as  
 CC diagnostic tests. The protein can also be used as a vaccine against  
 CC nematode parasites. (Updated on 06-AUG-2003 to correct OS field.)  
 XX  
 SQ Sequence 84 AA;

Query Match 100.0%; Score 486; DB 3; Length 84;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-38;  
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KATWQCGENKYSKSGKCDKCKYDGVVEEDDEEPNVPCLVRVCHQDCVCEGFYRNK 60  
 DB 1 KATWQCGENKYSKSGKCDKCKYDGVVEEDDEEPNVPCLVRVCHQDCVCEGFYRNK 60  
 QY 61 DDKVSAEDCELDNMDFIYPGTRN 84  
 DB 61 DDKVSAEDCELDNMDFIYPGTRN 84

RESULT 3  
 AAR91701  
 ID AAR91701 standard; protein; 91 AA.  
 XX  
 AC AAR91701;  
 DT 25-MAR-2003 (revised)  
 DT 17-NOV-1996 (first entry)  
 XX  
 DE AcanAPc2.  
 XX  
 KW AcanAP; HpoNAP; NamNAP; AcenAP; AduNAP; anticoagulant;  
 KW nematode-extracted anticoagulant protein; serine protease; nematode;  
 KW thrombosis; parasitic worm.  
 XX  
 OS Ancylostoma caninum.

XX WO9612021-A2.  
 XX 25-APR-1996.  
 XX 17-OCT-1995; 95WO-US013231.  
 XX 18-OCT-1994; 94US-00326110.  
 XX 05-JUN-1995; 95US-00461965.  
 XX 05-JUN-1995; 95US-00465380.  
 XX 05-JUN-1995; 95US-00486397.  
 XX 05-JUN-1995; 95US-00486399.  
 XX (CORV-) CORVAS INT INC.

XX Vlasuk GP, Stanssens PEH, Messens JHL, Lauwereys MJ, Laroche YR;  
 XX Jespers LS, Gansemans YGJ, Moyle M, Bergum PW;  
 XX WPI; 1996-222007/22.  
 XX N-PSDB; AAT12947.  
 XX Proteins with anticoagulant and/or serine protease inhibitory activity -  
 XX isolated from nematodes and useful to inhibit blood coagulation.

XX Claim 89 + 96; Fig 9; 243pp; English.  
 XX Proteins with anticoagulant and/or serine protease inhibitory activity,  
 XX isolated from nematodes, are useful to inhibit blood coagulation. The  
 XX proteins can be added to blood collection tubes defining the collection  
 XX of mammalian plasma. They are also useful to prevent or inhibit  
 XX thrombosis, and may be given alone or in combination with other  
 XX therapeutic or in vivo diagnostic agents. The proteins can serve as

CC immunogens to raise antibodies for use in the diagnosis and  
 CC identification of NAP concn. levels in biological fluids, e.g. to detect  
 CC mammalian infection with a parasitic worm. They can also be used as  
 CC immunogens in prophylactic and therapeutic vaccines against parasitic  
 CC worm infection. The proteins may double the clotting time of human plasma  
 CC in prothrombin time assays when present at 10-50 nMol, and double the  
 CC clotting time of human plasma in activated partial thrombin time assays  
 CC when present at 10-100 nMol. The anticoagulant proteins are pref. derived  
 CC from Ancylostoma caninum, A. ceylanicum, A. duodenale, Necator americanus  
 CC or Heligmosomoides polygyrus. The proteins pref. have 2 NAP domains and  
 CC specifically inhibit the catalytic activity of the factor VIIa/TF complex  
 CC in the presence of factor Xa or a catalytically inactive factor Xa  
 CC deriv., do not specifically inhibit the activation of factor VIIa in the  
 CC absence of TF and do not specifically inhibit prothrombinase. (Updated on  
 CC 25-MAR-2003 to correct PI field.)  
 XX

XX SQ Sequence 91 AA;

Query Match 100.0%; Score 486; DB 2; Length 91;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-38; Indels 0; Gaps 0;  
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATMQCGENEKDYSCGSKCEDKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFGYRNK 60  
 |||||  
 Db 8 KATMQCGENEKDYSCGSKCEDKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFGYRNK 67

QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84  
 |||||  
 Db 68 DDKCVSAEDCELDNMDFIYPGTRN 91

RESULT 4

AAV30393  
 ID AAY30393 standard; protein; 91 AA.

XX AC AAY30393;

XX DT 15-NOV-1999 (first entry)

XX DE Nematode extracted anticoagulant protein AcanAPc2.

XX KW Nematode extracted anticoagulant protein; NAP; anticoagulant;  
 KW serine protease inhibitor; NAP domain; factor VIIa/TF.

XX OS Ancylostoma caninum.

XX PN US5955294-A.

XX PD 21-SEP-1999.

XX PF 19-APR-1996; 96US-00634641.

XX PR 18-OCT-1994; 94US-00326110.

XX PR 05-JUN-1995; 95US-00461965.

XX PR 05-JUN-1995; 95US-00465380.

XX PR 05-JUN-1995; 95US-00486397.

XX PR 05-JUN-1995; 95US-00486399.

XX PR 17-OCT-1995; 95WO-US013231.

XX PA (CORV-) CORVAS INT INC.

XX PI Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;

PI Bergum FW, Messens JHL, Laroche YR, Vlasuk GP;

XX WPI; 1999-539569/45.

XX DR N-PSDB; AAZ99999.

XX PT Screening an isolated protein for Nematode-extracted Anticoagulant

PT Protein domains.

XX Example 13; Fig 9; 197pp; English.

XX The present sequence represents a nematode extracted anticoagulant

CC protein (NAP). The protein has activity as an anticoagulant and/or serine  
 CC protease inhibitor. The protein contains at least one NAP domain which  
 CC has selective inhibitory activity for factor VIIa/TF. The specification  
 CC describes a method for screening an isolated protein at least one domain  
 CC for factor VIIa/TF selective inhibitory activity. The method comprises  
 CC determining the time to clotting effected by a concentration of the  
 CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo  
 CC activated partial thromboplastin time (APTT) assay; calculating  
 CC prolongation of clotting effected by the isolated protein in each of the  
 CC PT and APTT assay, with respect to a baseline clotting value for each  
 CC assay, where prolongation of clotting is calculated as fold elevation of  
 CC clotting time relative to a baseline clotting value, where a doubling of  
 CC clotting time is deemed a two-fold elevation; and calculating a PT to  
 CC APTT prolongation ratio, where a ratio at least one is indicative of  
 CC factor VIIa/TF inhibitory activity. The method is useful for determining  
 CC if a protein has factor VIIa/TF inhibitory activity  
 XX

XX SQ Sequence 91 AA;

Query Match 100.0%; Score 486; DB 2; Length 91;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-38; Indels 0; Gaps 0;  
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATMQCGENEKDYSCGSKCEDKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFGYRNK 60  
 |||||  
 Db 8 KATMQCGENEKDYSCGSKCEDKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFGYRNK 67

QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84  
 |||||  
 Db 68 DDKCVSAEDCELDNMDFIYPGTRN 91

RESULT 5

AAV30454  
 ID AAY30454 standard; protein; 91 AA.

XX AC AAY30454;

XX DT 15-NOV-1999 (first entry)

XX DE Nematode extracted anticoagulant protein AcanAPc2.

XX KW Nematode extracted anticoagulant protein; NAP; anticoagulant;  
 KW serine protease inhibitor; NAP domain; factor VIIa/TF.

XX OS Ancylostoma caninum.

XX PN US5955294-A.

XX PD 21-SEP-1999.

XX PF 19-APR-1996; 96US-00634641.

XX PR 18-OCT-1994; 94US-00326110.

XX PR 05-JUN-1995; 95US-00461965.

XX PR 05-JUN-1995; 95US-00465380.

XX PR 05-JUN-1995; 95US-00486397.

XX PR 05-JUN-1995; 95US-00486399.

XX PR 17-OCT-1995; 95WO-US013231.

XX PA (CORV-) CORVAS INT INC.

XX PI Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;

PI Bergum FW, Messens JHL, Laroche YR, Vlasuk GP;

XX WPI; 1999-539569/45.

XX PT Screening an isolated protein for Nematode-extracted Anticoagulant

PT Protein domains.

XX Disclosure; Col 175-176; 197pp; English.

XX The present sequence represents a nematode extracted anticoagulant

CC protein (NAP). The protein has activity as an anticoagulant and/or serine  
 CC protease inhibitor. The protein contains at least one NAP domain which  
 CC has selective inhibitory activity for factor VIIa/TF. The specification  
 CC describes a method for screening an isolated protein at least one domain  
 CC for factor VIIa/TF selective inhibitory activity. The method comprises  
 CC determining the time to clotting effected by a concentration of the  
 CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo  
 CC activated partial thromboplastin time (APTT) assay; calculating  
 CC prolongation of clotting effected by the isolated protein in each of the  
 CC PT and APTT assay, with respect to a baseline clotting value for each  
 CC assay, where prolongation of clotting is calculated as fold elevation of  
 CC clotting time relative to a baseline clotting value, where a doubling of  
 CC clotting time is deemed a two-fold elevation; and calculating a PT to  
 CC APTT prolongation ratio, where a ratio at least one is indicative of  
 CC factor VIIa/TF inhibitory activity. The method is useful for determining  
 CC if a protein has factor VIIa/TF inhibitory activity  
 XX  
 SQ Sequence 91 AA;

Query Match 100.0%; Score 486; DB 2; Length 91;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-38;  
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATMQCGENKYSKSGKEDCKKCKYDGVVEEDDEEPNVPCLVRVCHQDCVCEGFRNK 60  
 DB 8 KATMQCGENKYSKSGKEDCKKCKYDGVVEEDDEEPNVPCLVRVCHQDCVCEGFRNK 57  
 QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84  
 DB 68 DDKCVSAEDCELDNMDFIYPGTRN 91

RESULT 6  
 AAB15346  
 ID AAB15346 standard; protein; 91 AA.  
 AC AAB15346;  
 DT 06-AUG-2003 (revised)  
 DT 19-DEC-2000 (first entry)  
 XX  
 DE A. caninum nematode-extracted anticoagulant protein AcanAPc2.  
 XX  
 KW Nematode-extracted anticoagulant protein; AcanAPc2; blood clotting;  
 KW canine hookworm; thrombosis; vaccine.  
 XX  
 OS Ancylostoma caninum.  
 XX  
 FN US6087487-A.  
 XX  
 PD 11-JUL-2000.  
 XX  
 PF 12-FEB-1999; 99US-00249451.  
 XX  
 PR 18-OCT-1994; 94US-00326110.  
 PR 05-JUN-1995; 95US-00461965.  
 PR 05-JUN-1995; 95US-00465380.  
 PR 05-JUN-1995; 95US-00486397.  
 PR 05-JUN-1995; 95US-00486399.  
 PR 17-OCT-1995; 95WO-US013231.  
 PR 17-APR-1997; 97US-00809455.  
 XX  
 PA (CORV-) CORVAS INT INC.  
 XX  
 PI Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;  
 PI Bergum PW, Messens JHL, Laroche YR, Vlasuk GP;  
 XX  
 DR WPI; 2000-531359/48.  
 DR N-PSDB; AAA73373.  
 XX  
 XX New cDNA molecule encoding a protein having factor Xa inhibitory activity  
 PT for preventing and treating blood clotting disorders, comprises nematode-  
 PT extracted anticoagulant protein domains.

XX Disclosure; Fig 11; 197pp; English.  
 PS  
 CC The present sequence is the Ancylostoma caninum nematode-extracted  
 CC anticoagulant protein AcanAPc2. Proteins of this kind have been shown to  
 CC be effective at preventing blood clotting without causing excessive  
 CC bleeding. The protein can be used in blood collection tubes to aid the  
 CC isolation of plasma from the blood, to prevent thrombosis which may be  
 CC linked to the rupture of an atherosclerotic plaque, acute myocardial  
 CC infarction, angina, thrombolytic therapy, percutaneous transluminal  
 CC coronary angioplasty, disseminated intravascular coagulopathy, infection,  
 CC cancer and septic shock, and to produce antibodies. In the latter  
 CC instance, the antibodies can be raised in order to detect infection by  
 CC nematodes (the coding sequence can also be used for this) or as  
 CC diagnostic tests. The protein can also be used as a vaccine against  
 CC nematode parasites. (Updated on 06-AUG-2003 to correct OS field.)  
 XX  
 SQ Sequence 91 AA;

Query Match 100.0%; Score 486; DB 3; Length 91;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-38;  
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATMQCGENKYSKSGKEDCKKCKYDGVVEEDDEEPNVPCLVRVCHQDCVCEGFRNK 60  
 DB 8 KATMQCGENKYSKSGKEDCKKCKYDGVVEEDDEEPNVPCLVRVCHQDCVCEGFRNK 67  
 QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84  
 DB 68 DDKCVSAEDCELDNMDFIYPGTRN 91

RESULT 7  
 AAY30422  
 ID AAY30422 standard; protein; 82 AA.  
 AC AAY30422;  
 DT 15-NOV-1999 (first entry)  
 XX  
 DE Mature nematode extracted anticoagulant protein AcanAPd2.  
 XX  
 KW Nematode extracted anticoagulant protein; NAP; anticoagulant;  
 KW serine protease inhibitor; NAP domain; factor VIIa/TF.  
 XX  
 OS Ancylostoma ceylanicum.  
 XX  
 FN US5955294-A.  
 XX  
 PD 21-SEP-1999.  
 XX  
 PF 19-APR-1996; 96US-00634641.  
 XX  
 PR 18-OCT-1994; 94US-00326110.  
 PR 05-JUN-1995; 95US-00461965.  
 PR 05-JUN-1995; 95US-00465380.  
 PR 05-JUN-1995; 95US-00486397.  
 PR 05-JUN-1995; 95US-00486399.  
 PR 17-OCT-1995; 95WO-US013231.  
 XX  
 PA (CORV-) CORVAS INT INC.  
 XX  
 PI Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;  
 PI Bergum PW, Messens JHL, Laroche YR, Vlasuk GP;  
 XX  
 DR WPI; 1999-539569/45.  
 XX  
 PT Screening an isolated protein for Nematode-extracted Anticoagulant  
 PT Protein Domains.  
 XX  
 PS Disclosure; Col 135-136; 197pp; English.  
 XX  
 CC The present sequence represents a nematode extracted anticoagulant



XX The present sequence is the Ancylostoma ceylanicum nematode-extracted  
CC anticoagulant protein AceNAP4d2. Proteins of this kind have been shown to  
CC be effective at preventing blood clotting without causing excessive  
CC bleeding. The protein can be used in blood collection tubes to aid the  
CC isolation of plasma from the blood, to prevent thrombosis which may be  
CC linked to the rupture of an atherosclerotic plaque, acute myocardial  
CC infarction, angina, thrombolytic therapy, percutaneous transluminal  
CC coronary angioplasty, disseminated intravascular coagulopathy, infection,  
CC cancer and septic shock, and to produce antibodies. In the latter  
CC instances, the antibodies can be raised in order to detect infection by  
CC nematodes or as diagnostic tests. The protein can also be used as a  
CC vaccine against nematode parasites. (Updated on 06-AUG-2003 to correct OS  
CC field.)  
XX  
XX  
SQ Sequence 82 AA;  
  
Query Match 43.6%; Score 212; DB 3; Length 82;  
Best Local Similarity 50.0%; Pred. No. 1.6e-12;  
Matches 38; Conservative 11; Mismatches 19; Indels 8; Gaps 3  
  
QY 6 CGENKYPSCGS-KECDKKCKYGVVEEDDEEPNVPCLVRVCHQD--CVCEGFYRNKDD 62  
DB 4 CGSNERYSDCCNDKQCKRKNEDDYKGDZ-----ACRSHVCEPACVCEGDFYRNKKG 58  
QY 63 KCVSAEDCELDNMDFI 78  
DB 59 SCVESDDCEYDNMDFI 74  
  
RESULT 11  
AAY30435  
ID AAY30435 standard; protein; 171 AA.  
AC AAY30435;  
XX  
DT 15-NOV-1999 (first entry) XX  
XX Mature nematode extracted anticoagulant protein AceNAP4. XX  
XX Nematode extracted anticoagulant protein; NAP; anticoagulant; XX  
XX serine protease inhibitor; NAP domain; factor VIIa/TF. XX  
XX Ancylostoma ceylanicum. XX  
XX  
XX US5955294-A. XX  
XX  
XX 21-SEP-1999. XX  
XX  
XX 19-APR-1996; 96US-00634641. XX  
XX  
XX 18-OCT-1994; 94US-00326110. XX  
XX 05-JUN-1995; 95US-00461965. XX  
XX 05-JUN-1995; 95US-00465380. XX  
XX 05-JUN-1995; 95US-00486397. XX  
XX 05-JUN-1995; 95US-00486399. XX  
XX 17-OCT-1995; 95WO-US013231. XX  
XX  
XX (CORV-) CORVAS INT INC. XX  
XX  
XX Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;  
PI Bergum PW, Messens JHL, Laroche YR, Vlasuk GP;  
XX  
XX WPI; 1999-539569/45. XX  
XX  
XX Screening an isolated protein for Nematode-extracted Anticoagulant  
XX Protein domains.  
XX  
XX Disclosure; Fig 17; 197pp; English. XX  
XX  
XX The present sequence represents a nematode extracted anticoagulant  
XX protein (NAP). The protein has activity as an anticoagulant and/or serine  
XX protease inhibitor. The protein contains at least one NAP domain which

CC has selective inhibitory activity for factor VIIa/TF. The specification  
 CC describes a method for screening an isolated protein at least one domain  
 CC for factor VIIa/TF selective inhibitory activity. The method comprises  
 CC determining the time to clotting effected by a concentration of the  
 CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo  
 CC activated partial thromboplastin time (APTT) assay; calculating  
 CC prolongation of clotting effected by the isolated protein in each of the  
 CC PT and APTT assay, with respect to a baseline clotting value for each  
 CC assay, where prolongation of clotting is calculated as fold elevation of  
 CC clotting time relative to a baseline clotting value, where a doubling of  
 CC clotting time is deemed a two-fold elevation; and calculating a PT to  
 CC APTT prolongation ratio, where a ratio at least one is indicative of  
 CC factor VIIa/TF inhibitory activity. The method is useful for determining  
 CC if a protein has factor VIIa/TF inhibitory activity  
 XX  
 SQ Sequence 171 AA;

Query Match 43.6%; Score 212; DB 2; Length 171;  
 Best Local Similarity 50.0%; Pred. No. 3.5e-12;  
 Matches 38; Conservative 11; Mismatches 19; Indels 8; Gaps 3;  
 QY 6 CGENKDYDCGS-KECDKCKYGVVEEDDEPNVPCLVVCHQD--CVCEGFGYRNKDD 62  
 DB 93 CGSNERYSDCGNDKQCKERNEDDYKGD-----ACRSHVCFPGACVCEGFGYRNKKG 147  
 QY 63 KCVSAEDCELDNMDFI 78  
 DB 148 SCVESDDCEYDNMDFI 163

RESULT 12  
 AAR91710  
 ID AAR91710 standard; protein; 190 AA.  
 AC AAR91710;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 17-NOV-1996 (first entry)  
 XX  
 DE ACeNAP4.  
 XX  
 KW ACaNaP; HpoNaP; NamNaP; AceNaP; AduNaP; antioagulant;  
 KW nematode-extracted antioagulant protein; serine protease; nematode;  
 KW thrombosis; parasitic worm.  
 XX  
 OS Ancylostoma ceylanicum.

Key Location/Qualifiers  
 FT Domain 1..108 /label= AceNaP4d1  
 FT Domain 109..190 /label= AceNaP4d2  
 FT  
 XX WO9612021-A2.  
 XX  
 XX 25-APR-1996.  
 XX  
 XX 17-OCT-1995; 95WO-US013231.  
 XX  
 XX 18-OCT-1994; 94US-00326110.  
 XX  
 XX 05-JUN-1995; 95US-00461965.  
 XX  
 XX 05-JUN-1995; 95US-00465380.  
 XX  
 XX 05-JUN-1995; 95US-00486397.  
 XX  
 XX (CORV-) CORVAS INT INC.

XX Vlausk GP, Stanssens PEH, Messens JHL, Lauwereys MJ, Laroche YR;  
 XX Jespers LS, Ganssema YGJ, Moyle M, Bergum PW;  
 XX MPI; 1996-222007/22.  
 XX N-PSDB; AAT12956.

PT Proteins with antioagulant and/or serine protease inhibitory activity -  
 PT isolated from nematodes and useful to inhibit blood coagulation.

XX Claim 221; Fig 7A; 243pp; English.

XX Proteins with antioagulant and/or serine protease inhibitory activity,  
 CC isolated from nematodes, are useful to inhibit blood coagulation. The  
 CC proteins can be added to blood collection tubes defining the collection  
 CC of mammalian plasma. They are also useful to prevent or inhibit  
 CC thrombosis, and may be given alone or in combination with other  
 CC therapeutic or in vivo diagnostic agents. The proteins can serve as  
 CC immunogens to raise antibodies for use in the diagnosis and  
 CC identification of NAP concn. levels in biological fluids, e.g. to detect  
 CC mammalian infection with a parasitic worm. They can also be used as  
 CC immunogens in prophylactic and therapeutic vaccines against parasitic  
 CC worm infection. The proteins may double the clotting time of human plasma  
 CC in prothrombin time assays when present at 10-50 nMol, and double the  
 CC clotting time of human plasma in activated partial thrombin time assays  
 CC when present at 10-100 nMol. The antioagulant proteins are pref. derived  
 CC from Ancylostoma caninum, A. ceylanicum, A. duodenale, Necator americanus  
 CC or Heligmosomoides polygyrus. The proteins pref. have 2 NAP domains and  
 CC specifically inhibit the catalytic activity of the factor VIIa/TF complex  
 CC in the presence of factor Xa or a catalytically inactive factor Xa  
 CC deriv., do not specifically inhibit the activation of factor VIIa in the  
 CC absence of TF and do not specifically inhibit prothrombinase. (Updated on  
 CC 25-MAR-2003 to correct PI field.)  
 XX

SQ Sequence 190 AA;

Query Match 43.6%; Score 212; DB 2; Length 190;  
 Best Local Similarity 50.0%; Pred. No. 3.9e-12;  
 Matches 38; Conservative 11; Mismatches 19; Indels 8; Gaps 3;

QY 6 CGENKDYDCGS-KECDKCKYGVVEEDDEPNVPCLVVCHQD--CVCEGFGYRNKDD 62  
 DB 112 CGSNERYSDCGNDKQCKERNEDDYKGD-----ACRSHVCFPGACVCEGFGYRNKKG 166  
 QY 63 KCVSAEDCELDNMDFI 78  
 DB 167 SCVESDDCEYDNMDFI 182

RESULT 13  
 AAY30384  
 ID AAY30384 standard; protein; 190 AA.  
 AC AAY30384;  
 XX  
 DT 15-NOV-1999 (first entry)  
 DT  
 XX  
 DE Nematode extracted antioagulant protein AceNaP4.  
 XX  
 KW Nematode extracted antioagulant protein; NAP; antioagulant;  
 KW serine protease inhibitor; NAP domain; factor VIIa/TF.  
 XX  
 OS Ancylostoma ceylanicum.

XX US5955294-A.  
 XX  
 XX 21-SEP-1999.  
 XX  
 XX 19-APR-1996; 96US-00634641.  
 XX  
 XX 18-OCT-1994; 94US-00326110.  
 XX  
 XX 05-JUN-1995; 95US-00461965.  
 XX  
 XX 05-JUN-1995; 95US-00465380.  
 XX  
 XX 05-JUN-1995; 95US-00486397.  
 XX  
 XX 05-JUN-1995; 95US-00486399.  
 XX  
 XX 17-OCT-1995; 95WO-US013231.  
 XX  
 XX (CORV-) CORVAS INT INC.

XX Lauwereys MJ, Stanssens PEH, Jespers LS, Ganssema YGJ, Moyle M;  
 XX

```

PI Bergum PW, Messens JHL, Laroche YR, Vlasuk GP;
XX WPI; 1999-539569/45.
DR N-PSDB; AAZ10441.
XX
XX Screening an isolated protein for Nematode-extracted Anticoagulant
PT Protein domains.
XX
XX Example 9; Fig 7A; 197pp; English.
XX
XX The present sequence represents a nematode extracted anticoagulant
CC protein (NAP). The protein has activity as an anticoagulant and/or serine
CC protease inhibitor. The protein contains at least one NAP domain which
CC has selective inhibitory activity for factor VIIa/TF. The specification
CC describes a method for screening an isolated protein at least one domain
CC for factor VIIa/TF selective inhibitory activity. The method comprises
CC determining the time to clotting effected by a concentration of the
CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo
CC activated partial thromboplastin time (aPTT) assay; calculating
CC prolongation of clotting effected by the isolated protein in each of the
CC PT and aPTT assay, with respect to a baseline clotting value for each
CC assay, where prolongation of clotting is calculated as fold elevation of
CC clotting time relative to a baseline clotting value, where a doubling of
CC clotting time is deemed a two-fold elevation; and calculating a PT to
CC clotting time ratio, where a ratio at least one is indicative of
CC aPTT prolongation ratio, where a ratio at least one is indicative of
CC factor VIIa/TF inhibitory activity. The method is useful for determining
CC if a protein has factor VIIa/TF inhibitory activity
XX
XX Sequence 190 AA;
SQ
Query Match 43.6%; Score 212; DB 2; Length 190;
Best Local Similarity 50.0%; Pred. No. 3.9e-12; Indels 8; Gaps 3;
Matches 38; Conservative 11; Mismatches 19;
QY 6 CGGENEKYDSCGS-KECDKKCKYDGVDEEDDEBNVPCLVRVCHQD--CYCEGEGFYRNKDD 62
DB 112 CGSNERYSDGNDKQCKERKNEDDYKGDGDE----ACRSHVCERPACVCEDEGFGFYRNKKG 166
QY 63 KCVSAEDCELDNDPFI 78
DB 167 SCVESDDCEYDNDPFI 182
XX
RESULT 14
ID AAY30419 standard; protein; 86 AA.
XX
XX AAY30419;
XX
XX 15-NOV-1999 (first entry)
XX
XX Mature nematode extracted anticoagulant protein AcanAP44.
XX
XX Nematode extracted anticoagulant protein; NAP; anticoagulant;
XX serine protease inhibitor; NAP domain; factor VIIa/TF.
XX
XX Ancylostoma caninum.
XX
XX US955294-A.
XX
XX 21-SEP-1999.
XX
XX 19-APR-1996; 96US-00634641.
XX
XX 18-OCT-1994; 94US-00326110.
XX
XX 05-JUN-1995; 95US-00461965.
XX
XX 05-JUN-1995; 95US-00465380.
XX
XX 05-JUN-1995; 95US-00486397.
XX
XX 05-JUN-1995; 95US-00486399.
XX
XX 17-OCT-1995; 95WO-US013231.
XX
XX (CORV-) CORVAS INT INC.
XX
PI Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;
PI Bergum PW, Messens JHL, Laroche YR, Vlasuk GP;
XX WPI; 1999-539569/45.
XX
XX Screening an isolated protein for Nematode-extracted Anticoagulant
PT Protein domains.
XX
XX Disclosure; Col 131-134; 197pp; English.
XX
XX The present sequence represents a nematode extracted anticoagulant
CC protein (NAP). The protein has activity as an anticoagulant and/or serine
CC protease inhibitor. The protein contains at least one NAP domain which
CC has selective inhibitory activity for factor VIIa/TF. The specification
CC describes a method for screening an isolated protein at least one domain
CC for factor VIIa/TF selective inhibitory activity. The method comprises
CC determining the time to clotting effected by a concentration of the
CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo
CC activated partial thromboplastin time (aPTT) assay; calculating
CC prolongation of clotting effected by the isolated protein in each of the
CC PT and aPTT assay, with respect to a baseline clotting value for each
CC assay, where prolongation of clotting is calculated as fold elevation of
CC clotting time relative to a baseline clotting value, where a doubling of
CC clotting time is deemed a two-fold elevation; and calculating a PT to
CC clotting time ratio, where a ratio at least one is indicative of
CC aPTT prolongation ratio, where a ratio at least one is indicative of
CC factor VIIa/TF inhibitory activity. The method is useful for determining
CC if a protein has factor VIIa/TF inhibitory activity
XX
XX Sequence 86 AA;
SQ
Query Match 41.6%; Score 202; DB 2; Length 86;
Best Local Similarity 47.4%; Pred. No. 1.5e-11;
Matches 37; Conservative 15; Mismatches 22; Indels 4; Gaps 2;
QY 5 QCGGENEKYDSCGS-KECDKKCKYDGVDEEDD--BEPNVPCLVRVCHQ--DCVCEGEGFYRNK 60
DB 5 KCGFGERLDCANKKPCPKCKIETSEEDDDVEETDVRCLRVYCVCEPFLKCKICKDGYRNK 64
QY 61 DDKCVSAEDCELDNDPFI 78
DB 65 KGECDVDDVQCEDFNEFI 82
XX
RESULT 15
ID AAB15304 standard; protein; 86 AA.
XX
XX AAB15304;
XX
XX 06-AUG-2003 (revised)
XX
XX 19-DEC-2000 (first entry)
XX
XX A. caninum nematode-extracted anticoagulant protein AcanAP44.
XX
XX Nematode-extracted anticoagulant protein; AcanAP44; blood clotting;
XX canine hookworm; thrombosis; vaccine.
XX
XX Ancylostoma caninum.
XX
XX US6087487-A.
XX
XX 11-JUL-2000.
XX
XX 12-FEB-1999; 99US-00249451.
XX
XX 18-OCT-1994; 94US-00326110.
XX
XX 05-JUN-1995; 95US-00461965.
XX
XX 05-JUN-1995; 95US-00465380.
XX
XX 05-JUN-1995; 95US-00486397.
XX
XX 05-JUN-1995; 95US-00486399.
XX
XX 17-OCT-1995; 95WO-US013231.
XX
XX 17-APR-1997; 97US-00809455.
XX

```



/// Page Blank (Aspic)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: September 24, 2004, 07:30:36 ; Search time 52.1684 Seconds  
(without alignments)  
83.127 Million cell updates/sec

Title: US-09-498-556C-59

Perfect score: 486

Sequence: 1 KATMOGENEKYDSCGKEC.....VSAEDCLNDMDFIYPGTRN 84

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/protdata/2/aa/5A COMB.pap.\*
- 2: /cgn2\_6/protdata/2/aa/5B COMB.pap.\*
- 3: /cgn2\_6/protdata/2/aa/6A COMB.pap.\*
- 4: /cgn2\_6/protdata/2/aa/6B COMB.pap.\*
- 5: /cgn2\_6/protdata/2/aa/PCTUS COMB.pap.\*
- 6: /cgn2\_6/protdata/2/aa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	486	100.0	84	2	US-08-465-380-59
2	486	100.0	84	2	US-08-486-397-59
3	486	100.0	84	2	US-08-486-397-59
4	486	100.0	84	2	US-08-486-397-59
5	486	100.0	84	2	US-08-461-965-59
6	486	100.0	84	2	US-08-634-641-59
7	486	100.0	84	3	US-09-249-471-59
8	486	100.0	84	3	US-09-249-471-59
9	486	100.0	84	3	US-09-249-451-59
10	486	100.0	84	3	US-08-809-455-59
11	486	100.0	84	3	US-09-249-461-59
12	486	100.0	84	3	US-09-249-448-59
13	486	100.0	84	4	US-09-249-473-59
14	486	100.0	91	2	US-08-465-380-128
15	486	100.0	91	2	US-08-480-478-50
16	486	100.0	91	2	US-08-486-397-128
17	486	100.0	91	2	US-08-486-399-128
18	486	100.0	91	2	US-08-461-965-128
19	486	100.0	91	2	US-08-326-110A-54
20	486	100.0	91	2	US-08-634-641-128
21	486	100.0	91	3	US-09-249-471-128
22	486	100.0	91	3	US-09-249-472-128
23	486	100.0	91	3	US-09-249-451-128
24	486	100.0	91	3	US-08-809-455-128
25	486	100.0	91	3	US-09-249-461-128
26	486	100.0	91	3	US-09-249-448-128
27	212	43.6	82	2	US-08-465-380-25

28	212	43.6	82	2	US-08-465-380-49
29	212	43.6	82	2	US-08-480-478-54
30	212	43.6	82	2	US-08-486-397-25
31	212	43.6	82	2	US-08-486-397-49
32	212	43.6	82	2	US-08-486-399-25
33	212	43.6	82	2	US-08-486-399-49
34	212	43.6	82	2	US-08-461-965-25
35	212	43.6	82	2	US-08-461-965-49
36	212	43.6	82	2	US-08-326-110A-54
37	212	43.6	82	2	US-08-634-641-25
38	212	43.6	82	2	US-08-634-641-49
39	212	43.6	82	3	US-09-249-471-25
40	212	43.6	82	3	US-09-249-471-49
41	212	43.6	82	3	US-09-249-472-25
42	212	43.6	82	3	US-09-249-472-49
43	212	43.6	82	3	US-09-249-451-25
44	212	43.6	82	3	US-09-249-451-49
45	212	43.6	82	3	US-08-809-455-25

## ALIGNMENTS

RESULT 1  
US-08-465-380-59  
; Sequence 59, Application US/08465380  
; Patent No. 5863894  
; GENERAL INFORMATION:  
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,  
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,  
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,  
; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,  
; APPLICANT: Peter W. Bergum  
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT  
; TITLE OF INVENTION: PROTEIN  
; NUMBER OF SEQUENCES: 356  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/465,380  
; FILING DATE: June 5, 1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/326,110  
; FILING DATE: October 18, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BIGGS, SUZANNE L.  
; REGISTRATION NUMBER: 30,158  
; REFERENCE/DOCKET NUMBER: 213/268  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 59:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 84 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; ORGANISM: Ancylostoma caninum

US-08-465-380-59

Query Match 100.0%; Score 486; DB 2; Length 84;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-43;  
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATWQCGENKDYSCGSKCKKCKYDGVVEEDDEBNPVCPLVRVCHQDCVCEGFRNK 60  
 |||  
 Db 1 KATWQCGENKDYSCGSKCKKCKYDGVVEEDDEBNPVCPLVRVCHQDCVCEGFRNK 60  
 |||

QY 61 DDKCVAEDCELDNMDFIYPGTRN 84  
 |||  
 Db 61 DDKCVAEDCELDNMDFIYPGTRN 84  
 |||

RESULT 2

US-08-486-397-59  
 ; Sequence 59, Application US/08486397  
 ; Patent No. 5866542  
 ; GENERAL INFORMATION:  
 ; APPLICANT: George P. Vlasuk, Patric H. Stanssens,  
 ; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,  
 ; APPLICANT: Yves R. Laroche, Laurent S. Jespers,  
 ; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,  
 ; APPLICANT: Peter W. Bergum  
 ; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT  
 ; TITLE OF INVENTION: PROTEIN  
 ; NUMBER OF SEQUENCES: 357  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Lyon & Lyon  
 ; STREET: 633 West Fifth Street  
 ; STREET: Suite 4700  
 ; CITY: Los Angeles  
 ; STATE: California  
 ; COUNTRY: U.S.A.  
 ; ZIP: 90071  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 ; MEDIUM TYPE: storage  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: IBM P.C. DOS 5.0  
 ; SOFTWARE: Word Perfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/486,397  
 ; FILING DATE: June 5, 1995  
 ; CLASSIFICATION: 530  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/326,110  
 ; FILING DATE: October 18, 1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: BIGGS, SUZANNE L.  
 ; REGISTRATION NUMBER: 30,158  
 ; REFERENCE/DOCKET NUMBER: 213/269  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (213) 489-1600  
 ; TELEFAX: (213) 955-0440  
 ; TELEX: 67-3510  
 ; INFORMATION FOR SEQ ID NO: 59:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 84 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Ancylostoma caninum  
 ; US-08-486-397-59

Query Match 100.0%; Score 486; DB 2; Length 84;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-43;  
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATWQCGENKDYSCGSKCKKCKYDGVVEEDDEBNPVCPLVRVCHQDCVCEGFRNK 60  
 |||

Db 1 KATWQCGENKDYSCGSKCKKCKYDGVVEEDDEBNPVCPLVRVCHQDCVCEGFRNK 60

QY 61 DDKCVAEDCELDNMDFIYPGTRN 84  
 |||  
 Db 61 DDKCVAEDCELDNMDFIYPGTRN 84  
 |||

RESULT 3

US-08-486-399-59  
 ; Sequence 59, Application US/08486399  
 ; Patent No. 5866543  
 ; GENERAL INFORMATION:  
 ; APPLICANT: George P. Vlasuk, Patric H. Stanssens,  
 ; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,  
 ; APPLICANT: Yves R. Laroche, Laurent S. Jespers,  
 ; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,  
 ; APPLICANT: Peter W. Bergum  
 ; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT  
 ; TITLE OF INVENTION: PROTEIN  
 ; NUMBER OF SEQUENCES: 356  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Lyon & Lyon  
 ; STREET: 633 West Fifth Street  
 ; STREET: Suite 4700  
 ; CITY: Los Angeles  
 ; STATE: California  
 ; COUNTRY: U.S.A.  
 ; ZIP: 90071  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 ; MEDIUM TYPE: storage  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: IBM P.C. DOS 5.0  
 ; SOFTWARE: Word Perfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/486,399  
 ; FILING DATE: June 5, 1995  
 ; CLASSIFICATION: 530  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/326,110  
 ; FILING DATE: October 18, 1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: BIGGS, SUZANNE L.  
 ; REGISTRATION NUMBER: 30,158  
 ; REFERENCE/DOCKET NUMBER: 213/270  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (213) 489-1600  
 ; TELEFAX: (213) 955-0440  
 ; TELEX: 67-3510  
 ; INFORMATION FOR SEQ ID NO: 59:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 84 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Ancylostoma caninum  
 ; US-08-486-399-59

Query Match 100.0%; Score 486; DB 2; Length 84;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-43;  
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATWQCGENKDYSCGSKCKKCKYDGVVEEDDEBNPVCPLVRVCHQDCVCEGFRNK 60  
 |||  
 Db 1 KATWQCGENKDYSCGSKCKKCKYDGVVEEDDEBNPVCPLVRVCHQDCVCEGFRNK 60  
 |||

QY 61 DDKCVAEDCELDNMDFIYPGTRN 84  
 |||  
 Db 61 DDKCVAEDCELDNMDFIYPGTRN 84  
 |||

RESULT 4

US-08-461-965-59  
; Sequence 59, Application US/08461965  
; Patent No. 5872098  
; GENERAL INFORMATION:  
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,  
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,  
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,  
; APPLICANT: Yannick G.J. Ganssemans, Matthew Moyle,  
; APPLICANT: Peter W. Bergum  
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT  
; NUMBER OF SEQUENCES: 356  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/461,965  
; FILING DATE: June 5, 1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/326,110  
; FILING DATE: October 18, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BIGGS, SUZANNE L.  
; REGISTRATION NUMBER: 30,158  
; REFERENCE/DOCKET NUMBER: 210/243  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 59:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 84 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; ORGANISM: Ancylostoma caninum  
; US-08-461-965-59  
  
Query Match 100.0%; Score 486; DB 2; Length 84;  
Best Local Similarity 100.0%; Pred. No. 4.7e-43;  
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 KATMQCGENKVDSCGSKCEDKCKYDGVVEEDDEEPNVPCLVRVCHQDCVCEGFRNK 60  
Db 1 KATMQCGENKVDSCGSKCEDKCKYDGVVEEDDEEPNVPCLVRVCHQDCVCEGFRNK 60  
  
QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84  
Db 61 DDKCVSAEDCELDNMDFIYPGTRN 84  
  
RESULT 5  
US-08-634-641-59  
; Sequence 59, Application US/08634641  
; Patent No. 5955294  
; GENERAL INFORMATION:  
; APPLICANT: Vlasuk, George P. Vlasuk  
; APPLICANT: Stanssens, Patrick Eric Hugo  
; APPLICANT: Mensens, Joris Hilda Lieven  
; APPLICANT: Lauwereys, Marc Josef

; APPLICANT: Laroche, Yves Rene  
; APPLICANT: Jespers, Laurent Stephane  
; APPLICANT: Ganssemans, Yannick Georges Jozef  
; APPLICANT: Moyle, Matthew  
; APPLICANT: Bergum, Peter W.  
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT  
; NUMBER OF SEQUENCES: 356  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/634,641  
; FILING DATE: April 19, 1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/13231  
; FILING DATE: October 17, 1995  
; APPLICATION NUMBER: 08/486,399  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/486,397  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/465,380  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/461,965  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/326,110  
; FILING DATE: October 18, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BIGGS, SUZANNE L.  
; REGISTRATION NUMBER: 30,158  
; REFERENCE/DOCKET NUMBER: 219/136  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 59:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 84 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; ORGANISM: Ancylostoma caninum  
; US-08-634-641-59  
  
Query Match 100.0%; Score 486; DB 2; Length 84;  
Best Local Similarity 100.0%; Pred. No. 4.7e-43;  
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 KATMQCGENKVDSCGSKCEDKCKYDGVVEEDDEEPNVPCLVRVCHQDCVCEGFRNK 60  
Db 1 KATMQCGENKVDSCGSKCEDKCKYDGVVEEDDEEPNVPCLVRVCHQDCVCEGFRNK 60  
  
QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84  
Db 61 DDKCVSAEDCELDNMDFIYPGTRN 84  
  
RESULT 6  
US-09-249-471-59  
; Sequence 59, Application US/09249471  
; Patent No. 6040441



Best Local Similarity 100.0%; Pred. No. 4.7e-43;  
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KATMOGGENEKYDSCGSKCEDKCKYDGVVEEDDEBNVPCLVRVCHQDCVCEGFGYRNK 60  
Db 1 KATMOGGENEKYDSCGSKCEDKCKYDGVVEEDDEBNVPCLVRVCHQDCVCEGFGYRNK 60

QY 61 DDKCVSAEDCELDNMDFIYFGTRN 84  
Db 61 DDKCVSAEDCELDNMDFIYFGTRN 84

RESULT 8

US-09-249-451-59  
; Sequence 59, Application US/09249451  
; Patent No. 6087487  
; GENERAL INFORMATION:  
; APPLICANT: Vlasuk, George Phillip  
; APPLICANT: Stanssens, Patrick Eric Hugo  
; APPLICANT: Messers, Joris Hilda Lieven  
; APPLICANT: Lauwereys, Marc Josef  
; APPLICANT: Laroche, Yves Rene  
; APPLICANT: Jespers, Laurent Stephane  
; APPLICANT: Gansemans, Yannick Georges Jozef  
; APPLICANT: Moyle, Matthew  
; APPLICANT: Bergum, Peter W.  
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
; TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT  
; TITLE OF INVENTION: PROTEIN  
; NUMBER OF SEQUENCES: 356  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071

COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/249,451  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/809,455  
; FILING DATE: April 17, 1997  
; APPLICATION NUMBER: PCT/US95/13231  
; FILING DATE: October 17, 1995  
; APPLICATION NUMBER: 08/486,399  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/486,397  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/465,380  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/461,965  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/326,110  
; FILING DATE: October 18, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BIGGS, SUZANNE L.  
; REGISTRATION NUMBER: 30,158  
; REFERENCE/DOCKET NUMBER: 216/270  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; INFORMATION FOR SEQ ID NO: 59:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 84 amino acids

; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; ORGANISM: Ancylostoma caninum  
US-09-249-451-59

Query Match 100.0%; Score 486; DB 3; Length 84;  
Best Local Similarity 100.0%; Pred. No. 4.7e-43;  
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATMOGGENEKYDSCGSKCEDKCKYDGVVEEDDEBNVPCLVRVCHQDCVCEGFGYRNK 60  
Db 1 KATMOGGENEKYDSCGSKCEDKCKYDGVVEEDDEBNVPCLVRVCHQDCVCEGFGYRNK 60  
QY 61 DDKCVSAEDCELDNMDFIYFGTRN 84  
Db 61 DDKCVSAEDCELDNMDFIYFGTRN 84

RESULT 9

US-08-809-455-59  
; Sequence 59, Application US/08809455  
; Patent No. 6090916  
; GENERAL INFORMATION:  
; APPLICANT: Vlasuk, George Phillip  
; APPLICANT: Stanssens, Patrick Eric Hugo  
; APPLICANT: Messers, Joris Hilda Lieven  
; APPLICANT: Lauwereys, Marc Josef  
; APPLICANT: Laroche, Yves Rene  
; APPLICANT: Jespers, Laurent Stephane  
; APPLICANT: Gansemans, Yannick Georges Jozef  
; APPLICANT: Moyle, Matthew  
; APPLICANT: Bergum, Peter W.  
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
; TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT  
; TITLE OF INVENTION: PROTEIN  
; NUMBER OF SEQUENCES: 356  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071

COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/809,455  
; FILING DATE: April 17, 1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/13231  
; FILING DATE: October 17, 1995  
; APPLICATION NUMBER: 08/486,399  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/486,397  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/465,380  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/461,965  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/326,110  
; FILING DATE: October 18, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BIGGS, SUZANNE L.  
; REGISTRATION NUMBER: 30,158  
; REFERENCE/DOCKET NUMBER: 216/270  
; TELECOMMUNICATION INFORMATION:

```
/ TELEPHONE: (213) 489-1600
/ TELEFAX: (213) 955-0440
/ TELEX: 67-3510
/ INFORMATION FOR SEQ ID NO: 59:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 84 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ ORIGINAL SOURCE:
/ ORGANISM: Ancylostoma caninum
/ US-08-809-455-59

Query Match 100.0%; Score 486; DB 3; Length 84;
Best Local Similarity 100.0%; Pred. No. 4.7e-43;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATMQCGENKDYSCGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEEGFYRNK 60
Db 1 KATMQCGENKDYSCGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEEGFYRNK 60

QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84
Db 61 DDKCVSAEDCELDNMDFIYPGTRN 84

RESULT 10
US-09-249-461-59
; Sequence 59, Application US/09249461
; Patent No. 6096877
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George Phillip
; APPLICANT: Stanssens, Patrick Eric Hugo
; APPLICANT: Messens, Joris Hilda Lieven
; APPLICANT: Lauwereys, Marc Josef
; APPLICANT: Laroche, Yves Rene
; APPLICANT: Jespers, Laurent Stephane
; APPLICANT: Gansemans, Yannick Georges Jozef
; APPLICANT: Moyle, Matthew
; APPLICANT: Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/249,461
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/809,455
; FILING DATE: April 17, 1997
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
```

```
/ FILING DATE: June 5, 1995
/ APPLICATION NUMBER: 08/326,110
/ FILING DATE: October 18, 1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BIGGS, SUZANNE L.
/ REGISTRATION NUMBER: 30,158
/ REFERENCE/DOCKET NUMBER: 216/270
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (213) 489-1600
/ TELEFAX: (213) 955-0440
/ TELEX: 67-3510
/ INFORMATION FOR SEQ ID NO: 59:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 84 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ ORIGINAL SOURCE:
/ ORGANISM: Ancylostoma caninum
/ US-09-249-461-59

Query Match 100.0%; Score 486; DB 3; Length 84;
Best Local Similarity 100.0%; Pred. No. 4.7e-43;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATMQCGENKDYSCGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEEGFYRNK 60
Db 1 KATMQCGENKDYSCGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEEGFYRNK 60

QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84
Db 61 DDKCVSAEDCELDNMDFIYPGTRN 84

RESULT 11
US-09-249-448-59
; Sequence 59, Application US/09249448
; Patent No. 6121435
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George Phillip
; APPLICANT: Stanssens, Patrick Eric Hugo
; APPLICANT: Messens, Joris Hilda Lieven
; APPLICANT: Lauwereys, Marc Josef
; APPLICANT: Laroche, Yves Rene
; APPLICANT: Jespers, Laurent Stephane
; APPLICANT: Gansemans, Yannick Georges Jozef
; APPLICANT: Moyle, Matthew
; APPLICANT: Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/249,448
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/809,455
; FILING DATE: April 17, 1997
; APPLICATION NUMBER: PCT/US95/13231
```

```
/ FILING DATE: October 17, 1995
/ APPLICATION NUMBER: 08/486,399
/ FILING DATE: June 5, 1995
/ APPLICATION NUMBER: 08/486,397
/ FILING DATE: June 5, 1995
/ APPLICATION NUMBER: 08/465,380
/ FILING DATE: June 5, 1995
/ APPLICATION NUMBER: 08/461,965
/ FILING DATE: June 5, 1995
/ APPLICATION NUMBER: 08/326,110
/ FILING DATE: October 18, 1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BIGGS, SUZANNE L.
/ REGISTRATION NUMBER: 30,158
/ REFERENCE/DOCKET NUMBER: 216/270
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (213) 489-1600
/ TELEFAX: (213) 955-0440
/ TELEX: 67-3510
/ INFORMATION FOR SEQ ID NO: 59:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 84 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ ORIGINAL SOURCE:
/ ORGANISM: Ancylostoma caninum
/
US-09-249-448-59

Query Match 100.0%; Score 486; DB 3; Length 84;
Best Local Similarity 100.0%; Pred. No. 4.7e-43;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATWCGGENEKYDSCGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFYRNK 60
Db 1 KATWCGGENEKYDSCGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFYRNK 60

QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84
Db 61 DDKCVSAEDCELDNMDFIYPGTRN 84

RESULT 12
US-09-249-473-59
/ Sequence 59, Application US/09249473
/ Patent No. 6534629
/ GENERAL INFORMATION:
/ APPLICANT: Vlasuk, George Phillip
/ APPLICANT: Stanssens, Patrick Eric Hugo
/ APPLICANT: Messens, Joris Hilda Lieven
/ APPLICANT: Lauwereys, Marc Josef
/ APPLICANT: Laroche, Yves Rene
/ APPLICANT: Jespers, Laurent Stephane
/ APPLICANT: Ganssemans, Yannick Georges Jozef
/ APPLICANT: Moyle, Matthew
/ APPLICANT: Bergum, Peter W.
/ TITLE OF INVENTION: NEMATOIDE-EXTRACTED SERINE PROTEASE
/ TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
/ TITLE OF INVENTION: PROTEIN
/ NUMBER OF SEQUENCES: 356
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Lyon & Lyon
/ STREET: 633 West Fifth Street
/ STREET: Suite 4700
/ CITY: Los Angeles
/ STATE: California
/ COUNTRY: U.S.A.
/ ZIP: 90071
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
/ MEDIUM TYPE: storage
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: IBM P.C. DOS 5.0
```

```
/ SOFTWARE: Word Perfect 5.1
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/249,473
/ FILING DATE:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/809,455
/ FILING DATE: April 17, 1997
/ APPLICATION NUMBER: PCT/US95/13231
/ FILING DATE: October 17, 1995
/ APPLICATION NUMBER: 08/486,399
/ FILING DATE: June 5, 1995
/ APPLICATION NUMBER: 08/486,397
/ FILING DATE: June 5, 1995
/ APPLICATION NUMBER: 08/465,380
/ FILING DATE: June 5, 1995
/ APPLICATION NUMBER: 08/461,965
/ FILING DATE: June 5, 1995
/ APPLICATION NUMBER: 08/326,110
/ FILING DATE: October 18, 1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BIGGS, SUZANNE L.
/ REGISTRATION NUMBER: 30,158
/ REFERENCE/DOCKET NUMBER: 216/270
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (213) 489-1600
/ TELEFAX: (213) 955-0440
/ TELEX: 67-3510
/ INFORMATION FOR SEQ ID NO: 59:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 84 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ ORIGINAL SOURCE:
/ ORGANISM: Ancylostoma caninum
/
US-09-249-473-59

Query Match 100.0%; Score 486; DB 4; Length 84;
Best Local Similarity 100.0%; Pred. No. 4.7e-43;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATWCGGENEKYDSCGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFYRNK 60
Db 1 KATWCGGENEKYDSCGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFYRNK 60

QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84
Db 61 DDKCVSAEDCELDNMDFIYPGTRN 84

RESULT 13
US-08-465-380-128
/ Sequence 128, Application US/08465380
/ Patent No. 5863894
/ GENERAL INFORMATION:
/ APPLICANT: George P. Vlasuk, Patric H. Stanssens,
/ APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
/ APPLICANT: Yves R. Laroche, Laurent S. Jespers,
/ APPLICANT: Yannick G.J. Ganssemans, Matthew Moyle,
/ APPLICANT: Peter W. Bergum
/ TITLE OF INVENTION: NEMATOIDE-EXTRACTED ANTICOAGULANT
/ TITLE OF INVENTION: PROTEIN
/ NUMBER OF SEQUENCES: 356
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Lyon & Lyon
/ STREET: 633 West Fifth Street
/ STREET: Suite 4700
/ CITY: Los Angeles
/ STATE: California
/ COUNTRY: U.S.A.
/ ZIP: 90071
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
```

MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/465,380  
FILING DATE: June 5, 1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 213/268  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 128:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 91 amino acids  
TYPE: amino acid  
MOLECULE TYPE: peptide  
TOPOLOGY: linear  
ORIGINAL SOURCE:  
ORGANISM: Ancylostoma caninum  
US-08-465-380-128

Query Match 100.0%; Score 486; DB 2; Length 91;  
Best Local Similarity 100.0%; Pred. No. 5.le-43;  
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATMQCGENKYSKSGKCDKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEEGFYRNK 60  
Db 8 KATMQCGENKYSKSGKCDKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEEGFYRNK 67

QY 61 DDKCVSAEDCELDNMDFIYPGTN 84  
Db 68 DDKCVSAEDCELDNMDFIYPGTN 91

RESULT 14  
US-08-480-478-50  
Sequence 50, Application US/08480478  
Patent No. 5864009  
GENERAL INFORMATION:  
APPLICANT: GEORGE P. VLASUK; PATRICK ERIC  
APPLICANT: HUGO STANSSENS; JORIS HILDA  
APPLICANT: LIEVEN MESSENS; MARC JOZEF  
APPLICANT: LAUWEREYS; YVES RENE LAROCHE;  
APPLICANT: LAURENT STEPHANE JESPEERS; and  
APPLICANT: YANNICK GEORGES JOZEF  
APPLICANT: GANSEMANS  
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTI-  
TITLE OF INVENTION: COAGULANT PROTEIN  
NUMBER OF SEQUENCES: 86  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/480,478

FILING DATE: 06-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/326,110  
FILING DATE: 18 OCTOBER 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 208/290  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 91 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-480-478-50

Query Match 100.0%; Score 486; DB 2; Length 91;  
Best Local Similarity 100.0%; Pred. No. 5.le-43;  
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATMQCGENKYSKSGKCDKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEEGFYRNK 60  
Db 8 KATMQCGENKYSKSGKCDKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEEGFYRNK 67

QY 61 DDKCVSAEDCELDNMDFIYPGTN 84  
Db 68 DDKCVSAEDCELDNMDFIYPGTN 91

RESULT 15  
US-08-486-397-128  
Sequence 128 Application US/08486397  
Patent No. 5866342  
GENERAL INFORMATION:  
APPLICANT: George P. Vlasuk, Patric H. Stanssens,  
APPLICANT: Joris H. Mensens, Marc J. Lauwereys,  
APPLICANT: Yves R. Laroche, Laurent S. Jespers,  
APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,  
APPLICANT: Peter W. Borzum  
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT  
TITLE OF INVENTION: PROTEIN  
NUMBER OF SEQUENCES: 357  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/486,397  
FILING DATE: June 5, 1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 213/269  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600

```

; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 128:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 91 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
US-08-486-397-128

Query Match      100.0%; Score 486; DB 2; Length 91;
Best Local Similarity 100.0%; Pred. No. 5.1e-43;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATMQGENEKYDSCGSKCDKCKYDGVVEEDDEEPNVPCLVRVCHQDCVCEGFYRNK 60
   |||||||
Db 8 KATMQGENEKYDSCGSKCDKCKYDGVVEEDDEEPNVPCLVRVCHQDCVCEGFYRNK 67
   |||||||

QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84
   |||||||
Db 68 DDKCVSAEDCELDNMDFIYPGTRN 91
   |||||||

```

Search completed: September 24, 2004, 07:50:00  
Job time : 53.1684 secs

This Page Blank (uspio)

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 07:35:06 ; Search time 723.284 Seconds  
(without alignments)  
37.345 Million cell updates/sec

Title: US-09-498-556C-59

Perfect score: 486

Sequence: 1 KATMOCGENEYKSCSKEC.....VSAEDCLNDMDFIYPTGRN 84

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 32158718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	486	100.0	84	10	US-09-498-272-59
2	486	100.0	91	10	US-09-498-272-128
3	212	43.6	82	10	US-09-498-272-25
4	212	43.6	82	10	US-09-498-272-49
5	212	43.6	171	10	US-09-498-272-62
6	202	41.6	86	10	US-09-498-272-46
7	201	41.4	89	10	US-09-498-272-48
8	201	41.4	108	10	US-09-498-272-24
9	199	40.9	86	10	US-09-498-272-45
10	197.5	40.6	88	10	US-09-498-272-43
11	195.5	40.2	84	10	US-09-498-272-63
12	195.5	40.2	162	10	US-09-498-272-51
13	193.5	39.8	84	10	US-09-498-272-64
14	193.5	39.8	162	10	US-09-498-272-52
15	186	38.3	83	10	US-09-498-272-52

16	186	38.3	102	10	US-09-498-272-27	Sequence 27, Appl
17	186	38.3	161	10	US-09-498-272-65	Sequence 65, Appl
18	184.5	38.0	87	10	US-09-498-272-44	Sequence 44, Appl
19	178	36.6	78	10	US-09-498-272-47	Sequence 47, Appl
20	159	32.7	77	10	US-09-498-272-58	Sequence 58, Appl
21	159	32.7	96	10	US-09-498-272-23	Sequence 23, Appl
22	140.5	28.9	149	12	US-10-467-020-26	Sequence 26, Appl
23	140.5	28.9	156	12	US-10-467-020-29	Sequence 29, Appl
24	139.5	28.7	75	10	US-09-498-272-6	Sequence 6, Appl
25	139.5	28.7	75	10	US-09-498-272-41	Sequence 41, Appl
26	139.5	28.7	73	10	US-09-498-272-8	Sequence 8, Appl
27	139.5	28.7	98	10	US-09-498-272-21	Sequence 21, Appl
28	137.5	28.3	76	12	US-10-467-020-28	Sequence 28, Appl
29	137.5	28.3	77	10	US-09-498-272-4	Sequence 4, Appl
30	137.5	28.3	77	10	US-09-498-272-40	Sequence 40, Appl
31	137.5	28.3	78	10	US-09-498-272-28	Sequence 28, Appl
32	137.5	28.3	78	10	US-09-498-272-56	Sequence 56, Appl
33	137.5	28.3	81	10	US-09-498-272-7	Sequence 7, Appl
34	137.5	28.3	100	10	US-09-498-272-20	Sequence 20, Appl
35	137	28.2	75	10	US-09-498-272-57	Sequence 57, Appl
36	136.5	28.1	94	10	US-09-498-272-22	Sequence 22, Appl
37	136.5	28.1	78	10	US-09-498-272-53	Sequence 53, Appl
38	136	28.0	75	10	US-09-498-272-26	Sequence 26, Appl
39	134.5	27.7	77	10	US-09-498-272-55	Sequence 55, Appl
40	134.5	27.7	78	10	US-09-498-272-54	Sequence 54, Appl
41	118	24.3	67	14	US-10-087-887-40	Sequence 40, Appl
42	118	24.3	67	16	US-10-038-854-185	Sequence 185, Appl
43	111	22.8	74	10	US-09-498-272-42	Sequence 42, Appl
44	103	21.2	2601	12	US-10-016-248-59	Sequence 59, Appl
45	103	21.2	2601	15	US-10-028-248A-76	Sequence 76, Appl

#### ALIGNMENTS

RESULT 1  
US-09-498-272-59  
Sequence 59, Application US/09498272  
Publication No. US2003013890A1  
GENERAL INFORMATION:

APPLICANT: Vlaauk, George Phillip  
Stanssens, Patrick Eric Hugo  
Messens, Joris Hilda Lieven  
Lauwereys, Marc Josef  
Laroche, Yves Rene  
Jespers, Laurent Stephane  
Ganseman, Yannick Georges Jozef  
Moyle, Matthew  
Berghum, Peter W.

TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE INHIBITORS AND ANTICOAGULANT PROTEIN

NUMBER OF SEQUENCES: 356

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street

Suite 4700

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071

COMPUTER READABLE FORM: 3.5" Diskette, 1.44 Mb

storage

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/498,272

FILING DATE: 04-Feb-2000

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/13231

FILING DATE: October 17, 1995

```

; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; SEQUENCE DESCRIPTION: SEQ ID NO: 59:
US-09-498-272-59

Query Match 100.0%; Score 486; DB 10; Length 84;
Best Local Similarity 100.0%; Pred. No. 1.13e-41;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATMQCGENEKYDCSGSKEDKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEEGFYRNK 60
Db 1 KATMQCGENEKYDCSGSKEDKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEEGFYRNK 60

QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84
Db 61 DDKCVSAEDCELDNMDFIYPGTRN 84

RESULT 2
US-09-498-272-128
; Sequence 128, Application US/09498272
; Publication No. US20030113890A1
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George Phillip
; Stanssens, Patrick Eric Hugo
; Messens, Joris Hilda Lieven
; Lauwereys, Marc Josef
; Laroche, Yves Rene
; Jespers, Laurent Stephane
; Gansemans, Yannick Georges Jozef
; Moyle, Matthew
; Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; INHIBITORS AND ANTICOAGULANT
; PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0

```

```

; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/498,272
; FILING DATE: 04-Feb-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 128:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 91 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; SEQUENCE DESCRIPTION: SEQ ID NO: 128:
US-09-498-272-128

Query Match 100.0%; Score 486; DB 10; Length 91;
Best Local Similarity 100.0%; Pred. No. 1.4e-41;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATMQCGENEKYDCSGSKEDKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEEGFYRNK 60
Db 8 KATMQCGENEKYDCSGSKEDKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEEGFYRNK 67

QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84
Db 68 DDKCVSAEDCELDNMDFIYPGTRN 91

RESULT 3
US-09-498-272-25
; Sequence 25, Application US/09498272
; Publication No. US20030113890A1
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George Phillip
; Stanssens, Patrick Eric Hugo
; Messens, Joris Hilda Lieven
; Lauwereys, Marc Josef
; Laroche, Yves Rene
; Jespers, Laurent Stephane
; Gansemans, Yannick Georges Jozef
; Moyle, Matthew
; Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; INHIBITORS AND ANTICOAGULANT
; PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California

```

COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/498,272  
FILING DATE: 04-Feb-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486,397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 82 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Ancylostoma ceylanicum  
SEQUENCE DESCRIPTION: SEQ ID NO: 25:  
US-09-498-272-25  
Query Match 43.6%; Score 212; DB 10; Length 82;  
Best Local Similarity 50.0%; Pred. No. 5.3e-14;  
Matches 38; Conservative 11; Mismatches 19; Indels 8; Gaps 3;  
QY 6 CGENEKYDCGS-KECDKCKYDGVVEEDDEPNVCLRVCHQD--CVCEEGFYRNKDD 62  
Db 4 CGSNERYSDCGNDKQCKERNEDDYKGD-----ACRSHVCEPFGACVCEGDFYRNKKG 58  
QY 63 KCVSAEDCELDNMDFI 78  
Db 59 SCVESDDCEYDNMDFI 74  
RESULT 4  
US-09-498-272-49  
; Sequence 49, Application US/09498272  
; Publication No. US20030113890A1  
; GENERAL INFORMATION:  
; APPLICANT: Vlasuk, George Phillip  
; Stanssens, Patrick Eric Hugo  
; Messens, Joris Hilda Lieven  
; Lauwereys, Marc Josef  
; Laroche, Yves Rene  
; Jespers, Laurent Stephane  
; Gansemans, Yannick Georges Jozef  
; Moyle, Matthew  
; Bergum, Peter W.  
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
; INHIBITORS AND ANTICOAGULANT  
; PROTEIN

NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/498,272  
FILING DATE: 04-Feb-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486,397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 82 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Ancylostoma ceylanicum  
SEQUENCE DESCRIPTION: SEQ ID NO: 49:  
US-09-498-272-49  
Query Match 43.6%; Score 212; DB 10; Length 82;  
Best Local Similarity 50.0%; Pred. No. 5.3e-14;  
Matches 38; Conservative 11; Mismatches 19; Indels 8; Gaps 3;  
QY 6 CGENEKYDCGS-KECDKCKYDGVVEEDDEPNVCLRVCHQD--CVCEEGFYRNKDD 62  
Db 4 CGSNERYSDCGNDKQCKERNEDDYKGD-----ACRSHVCEPFGACVCEGDFYRNKKG 58  
QY 63 KCVSAEDCELDNMDFI 78  
Db 59 SCVESDDCEYDNMDFI 74  
RESULT 5  
US-09-498-272-62  
; Sequence 62, Application US/09498272  
; Publication No. US20030113890A1  
; GENERAL INFORMATION:  
; APPLICANT: Vlasuk, George Phillip  
; Stanssens, Patrick Eric Hugo  
; Messens, Joris Hilda Lieven  
; Lauwereys, Marc Josef  
; Laroche, Yves Rene

Jespers, Laurent Stephane  
Gansmans, Yannick Georges Jozef  
Moyle, Matthew  
Bergum, Peter W.  
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
INHIBITORS AND ANTICOAGULANT  
PROTEIN

NUMBER OF SEQUENCES: 356

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
SUITE: Suite 4700

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071

COMPUTER READABLE FORM: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/498,272

FILING DATE: 04-Feb-2000

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/13231

FILING DATE: October 17, 1995

APPLICATION NUMBER: 08/486,399

FILING DATE: June 5, 1995

APPLICATION NUMBER: 08/486,397

FILING DATE: June 5, 1995

APPLICATION NUMBER: 08/465,380

FILING DATE: June 5, 1995

APPLICATION NUMBER: 08/461,965

FILING DATE: June 5, 1995

APPLICATION NUMBER: 08/326,110

FILING DATE: October 18, 1994

ATTORNEY/AGENT INFORMATION:

NAME: BIGGS, SUZANNE L.

REGISTRATION NUMBER: 30,158

REFERENCE/DOCKET NUMBER: 216/270

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 62:

SEQUENCE CHARACTERISTICS:

LENGTH: 171 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE:

ORGANISM: Ancylostoma ceylanicum

SEQUENCE DESCRIPTION: SEQ ID NO: 62:

US-09-498-272-62

Query Match 43.6%; Score 212; DB 10; Length 171;  
Best Local Similarity 50.0%; Pred. No. 1.1e-13;  
Matches 38; Conservative 11; Mismatches 19; Indels 8; Gaps 3;

QY 6 CGENKIDSCGS-KCKDKCKYDGVVEEDDEPNVPCLVVCHQD--CVCSEGFYRNKDD 62

DB 93 CGSNERYDCGNKQCKERKCNEDYEKGDE-----ACRSHVCEPFGACVCDGFGYRNKG 147

QY 63 KCVSAEDCELDNMDFI 78

DB 148 SCVESDDCEYDNDMDFI 163

RESULT 6

US-09-498-272-46

; Sequence 46, Application US/09498272

Publication No. US20030113890A1  
GENERAL INFORMATION:  
APPLICANT: Vlasuk, George Phillip  
Stanssens, Patrick Eric Hugo  
Messens, Joris Hilda Lieven  
Lauwereys, Marc Josef  
Laroche, Yves Rene  
Jespers, Laurent Stephane  
Gansmans, Yannick Georges Jozef  
Moyle, Matthew  
Bergum, Peter W.  
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
INHIBITORS AND ANTICOAGULANT  
PROTEIN

NUMBER OF SEQUENCES: 356

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
SUITE: Suite 4700

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071

COMPUTER READABLE FORM: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/498,272

FILING DATE: 04-Feb-2000

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/13231

FILING DATE: October 17, 1995

APPLICATION NUMBER: 08/486,399

FILING DATE: June 5, 1995

APPLICATION NUMBER: 08/486,397

FILING DATE: June 5, 1995

APPLICATION NUMBER: 08/465,380

FILING DATE: June 5, 1995

APPLICATION NUMBER: 08/461,965

FILING DATE: June 5, 1995

APPLICATION NUMBER: 08/326,110

FILING DATE: October 18, 1994

ATTORNEY/AGENT INFORMATION:

NAME: BIGGS, SUZANNE L.

REGISTRATION NUMBER: 30,158

REFERENCE/DOCKET NUMBER: 216/270

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 46:

SEQUENCE CHARACTERISTICS:

LENGTH: 86 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE:

ORGANISM: Ancylostoma caninum

SEQUENCE DESCRIPTION: SEQ ID NO: 46:

US-09-498-272-46

Query Match 41.6%; Score 202; DB 10; Length 86;  
Best Local Similarity 47.4%; Pred. No. 5.6e-13;  
Matches 37; Conservative 15; Mismatches 22; Indels 4; Gaps 2;

QY 5 QCGENKIDSCGS-KCKDKCKYDGVVEEDD--EENVPCLVCHQ--DCVCEGFYRNK 60

DB 5 KCGFGERLDCANKKPCPKCKIETSEEDDDVEEDVCLVRCVCEPFLCKICKDGYRNK 64

QY 61 DDKCVSAEDCELDNMDFI 78

Matches 38; Conservative 13; Mismatches 24; Indels 4; Gaps 3;

Qy 3 TMOGENEKYDSCGS-KECDKKCKYDGVBEEDDEENVPCLVCHOD--CVCEEGFYRN 59  
 Db 7 TNACGLNEYFACGCMKECHRON-EEENEERDEERITACLIIVCFRPGACVCKDGFYRN 65

Qy 60 KDDKCVSAEDCELDNMDFI 78  
 Db 66 RTGSCVEEDDCEYENMEFI 84

RESULT 8  
 US-09-498-272-48  
 ; Sequence 45, Application US/09498272  
 ; Publication No. US20030113890A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Vlasuk, George Phillip  
 ; Stanssens, Patrick Eric Hugo  
 ; Messens, Joris Hilda Lieven  
 ; Lauwereys, Marc Josef  
 ; Laroche, Yves Rene  
 ; Jespers, Laurent Stephane  
 ; Gansmans, Yannick Georges Jozef  
 ; Moyle, Matthew  
 ; Bergum, Peter W.  
 ; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
 ; INHIBITORS AND ANTICOAGULANT  
 ; PROTEIN  
 ; NUMBER OF SEQUENCES: 356  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Lyon & Lyon  
 ; STREET: 633 West Fifth Street  
 ; Suite 4700  
 ; CITY: Los Angeles  
 ; STATE: California  
 ; COUNTRY: U.S.A.  
 ; ZIP: 90071  
 ; COMPUTER READABLE FORM: 3.5" Diskette, 1.44 Mb  
 ; MEDIUM TYPE: storage  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: IBM P.C. DOS 5.0  
 ; SOFTWARE: Word Perfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/498,272  
 ; FILING DATE: 04-Feb-2000  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US95/13231  
 ; FILING DATE: October 17, 1995  
 ; APPLICATION NUMBER: 08/486,399  
 ; FILING DATE: June 5, 1995  
 ; APPLICATION NUMBER: 08/461,965  
 ; FILING DATE: June 5, 1995  
 ; APPLICATION NUMBER: 08/326,110  
 ; FILING DATE: October 18, 1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: BIGGS, SUZANNE L.  
 ; REGISTRATION NUMBER: 30,158  
 ; REFERENCE/DOCKET NUMBER: 216/270  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (213) 489-1600  
 ; TELEFAX: (213) 955-0440  
 ; TELE: 67-3510  
 ; INFORMATION FOR SEQ ID NO: 48:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 89 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Ancylostoma ceylanicum  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 48:  
 ; US-09-498-272-48  
 ; Query Match 41.4%; Score 201; DB 10; Length 89;  
 ; Best Local Similarity 48.1%; Pred. No. 7.4e-13;

Db 65 KGECVTDDVQCQEDFMEFI 82

RESULT 7  
 US-09-498-272-48  
 ; Sequence 45, Application US/09498272  
 ; Publication No. US20030113890A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Vlasuk, George Phillip  
 ; Stanssens, Patrick Eric Hugo  
 ; Messens, Joris Hilda Lieven  
 ; Lauwereys, Marc Josef  
 ; Laroche, Yves Rene  
 ; Jespers, Laurent Stephane  
 ; Gansmans, Yannick Georges Jozef  
 ; Moyle, Matthew  
 ; Bergum, Peter W.  
 ; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
 ; INHIBITORS AND ANTICOAGULANT  
 ; PROTEIN  
 ; NUMBER OF SEQUENCES: 356  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Lyon & Lyon  
 ; STREET: 633 West Fifth Street  
 ; Suite 4700  
 ; CITY: Los Angeles  
 ; STATE: California  
 ; COUNTRY: U.S.A.  
 ; ZIP: 90071  
 ; COMPUTER READABLE FORM: 3.5" Diskette, 1.44 Mb  
 ; MEDIUM TYPE: storage  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: IBM P.C. DOS 5.0  
 ; SOFTWARE: Word Perfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/498,272  
 ; FILING DATE: 04-Feb-2000  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US95/13231  
 ; FILING DATE: October 17, 1995  
 ; APPLICATION NUMBER: 08/486,399  
 ; FILING DATE: June 5, 1995  
 ; APPLICATION NUMBER: 08/461,965  
 ; FILING DATE: June 5, 1995  
 ; APPLICATION NUMBER: 08/326,110  
 ; FILING DATE: October 18, 1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: BIGGS, SUZANNE L.  
 ; REGISTRATION NUMBER: 30,158  
 ; REFERENCE/DOCKET NUMBER: 216/270  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (213) 489-1600  
 ; TELEFAX: (213) 955-0440  
 ; TELE: 67-3510  
 ; INFORMATION FOR SEQ ID NO: 48:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 89 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Ancylostoma ceylanicum  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 48:  
 ; US-09-498-272-48  
 ; Query Match 41.4%; Score 201; DB 10; Length 89;  
 ; Best Local Similarity 48.1%; Pred. No. 7.4e-13;

```

; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma ceylanicum
; SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-09-498-272-24

Query Match          41.4%; Score 201; DB 10; Length 108;
Best Local Similarity 48.1%; Pred. NO. 9e-13;
Matches 38; Conservative 13; Mismatches 24; Indels 4; Gaps 3;

QY 3 TMOGNEKYDCSGS-KECDKKCKYGVVEEDDEPNVPCLVRVCHQD--CVCEEGFYRN 59
Db 26 TNACGLNEVFACGNMECHERCN-BEENHERDERITACLRVCFRPGACVCKDGFYRN 84

QY 60 KDDKCVSAEDCELDNMDFI 78
Db 85 RTGSCVEEDDCEYENMEFI 103

RESULT 9
US-09-498-272-45
; Sequence 45, Application US/09498272
; Publication No. US20030113890A1
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George Phillip
; Stanssens, Patrick Eric Hugo
; Messens, Joris Hilda Lieven
; Lauwereys, Marc Josef
; Laroche, Yves Rene
; Jespers, Laurent Stephane
; Gansemans, Yannick Georges Jozef
; Moyle, Matthew
; Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; INHIBITORS AND ANTICOAGULANT
; PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/498,272
; FILING DATE: 04-Feb-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440

```

```

; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-09-498-272-45

Query Match          40.9%; Score 199; DB 10; Length 86;
Best Local Similarity 46.2%; Pred. NO. 1.1e-12;
Matches 36; Conservative 16; Mismatches 22; Indels 4; Gaps 2;

QY 5 QCGENEKYDCSGS-KECDKKCKYGVVEEDD--EENVPCLVRVCHQ--DCVCEEGFYRNK 60
Db 5 KCGPGERLDCANKKPCPEKCKIETSEEDDDVEDTDRCLVRVCEPLKCKDGYYRNK 64

QY 61 DDKCVSAEDCELDNMDFI 78
Db 65 KGECVTDVCOEDFMEFI 82

RESULT 10
US-09-498-272-43
; Sequence 43, Application US/09498272
; Publication No. US20030113890A1
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George Phillip
; Stanssens, Patrick Eric Hugo
; Messens, Joris Hilda Lieven
; Lauwereys, Marc Josef
; Laroche, Yves Rene
; Jespers, Laurent Stephane
; Gansemans, Yannick Georges Jozef
; Moyle, Matthew
; Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; INHIBITORS AND ANTICOAGULANT
; PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/498,272
; FILING DATE: 04-Feb-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994

```



MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/498,272  
FILING DATE: 04-Feb-2000

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486,397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994

ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 499-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 51:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 84 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

ORIGINAL SOURCE:  
ORGANISM: Ancylostoma caninum  
SEQUENCE DESCRIPTION: SEQ ID NO: 51:

US-09-498-272-51

Query Match 39.8%; Score 193.5; DB 10; Length 84;  
Best Local Similarity 50.0%; Pred. No. 4e-12;  
Matches 40; Conservative 11; Mismatches 26; Indels 3; Gaps 3;

QY 1 KATVCCGENKYSOCS-RECKKKKYDGVBEEDDEENVECLIVC-HODCVCEGFYR 58  
DB : :: ||||| :: -:: ||||| :: ||||| :: |:: |||||  
1 KSAAKGGLNFKLD-CGNLKAACKKCSDLNDEEYGEDESKRSRCIGRVCDGFGYR 59  
QY 59 NKDDKVCVSABDCELDNNDFI 78  
DB 60 NKKGCQVTDRDDCEYDNWEII 79

RESULT 14  
US-09-498-272-64  
Sequence 64, Application US/09498272  
Publication No. US20030113890A1

GENERAL INFORMATION:  
APPLICANT: Vlasuk, George Phillip  
Stanssens, Patrick Eric Hugo  
Messens, Joris Hilda Lieven  
Lauwereys, Marc Josef  
Laroche, Yves Rene  
Jaspers, Laurent Stephane  
Gansemans, Yannick Georges Jozef  
Moyle, Matthew  
Bergum, Peter W.

TITLE OF INVENTION: MEMTODE-EXTRACTED SERINE PROTEASE  
INHIBITORS AND ANTICOAGULANT

NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street  
CITY: Suite 4700  
STATE: Los Angeles  
COUNTRY: California  
ZIP: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/498,272  
FILING DATE: 04-Feb-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486,397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 64:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 162 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Ancylostoma caninum  
SEQUENCE DESCRIPTION: SEQ ID NO: 64:  
US-09-498-272-64

Query Match 39.8%; Score 193.5; DB 10; Length 162;  
Best Local Similarity 50.0%; Pred. No. 7.8e-12;  
Matches 40; Conservative 11; Mismatches 26; Indels 3; Gaps 3;  
QY 1 KATWCCGENEKYDSCGS-KECDKKCKYDGVVEEDDEPNVPLRVYV-HODQVCBERGFYR 58  
Db 1 KSAKKGUNELKD-CGNLKCCEKCSLDNEEDYCEDESKCRSRECIGRVVCDEGFYR 59  
QY 59 NKDDKCVSAEDCELDNMDFI 78  
Db 60 NKGKGVTRDDCEYDNMEII 79

## RESULT 15

US-09-498-272-52

Sequence 52, Application US/09498272

Publication No. US20030113890A1

GENERAL INFORMATION:

APPLICANT: Vlasuk, George Phillip

Stanssens, Patrick Eric Hugo

Messens, Joris Hilda Lieven

Laureys, Marc Josef

Laroche, Yves Rene

Jespers, Laurent Stephane

Ganssemaers, Yannick Georges Jozef

Moyle, Matthew

Bergum, Peter W.  
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
INHIBITORS AND ANTICOAGULANT  
PROTEIN  
NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/498,272  
FILING DATE: 04-Feb-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486,397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 52:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 83 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Ancylostoma duodenale  
SEQUENCE DESCRIPTION: SEQ ID NO: 52:  
US-09-498-272-52  
Query Match 38.3%; Score 186; DB 10; Length 83;  
Best Local Similarity 50.0%; Pred. No. 2.2e-11;  
Matches 40; Conservative 11; Mismatches 25; Indels 4; Gaps 4;  
QY 1 KATWCCGENEKYDSCGS-KECDKKCKYDGVVEEDDEPNVPLRVYV-HODQVCBERGFYR 58  
Db 1 KAAKKGUNELKD-CGNLKCCEKCS-DLESEYEDESKCRSRECISPRVCVDEGFYR 58  
QY 59 NKDDKCVSAEDCELDNMDFI 78  
Db 59 NKGKGVTRDDCEYDNMEII 78  
Search completed: September 24, 2004, 08:10:03  
Job time : 723.284 secs

This Page Blank (usp10)

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	125	25.7	561	2	T27318	hypothetical prote
2	125	25.7	626	2	T27319	hypothetical prote
3	113	23.3	490	2	T32003	hypothetical prote
4	112	23.0	249	2	T24604	hypothetical prote
5	109.5	22.5	1642	2	T19130	hypothetical prote
6	105.5	21.7	1445	2	T15608	hypothetical prote
7	100	20.6	5376	2	T42215	zonadhesin - mouse
8	99	20.4	1036	2	T117405	scavenger receptor
9	98.5	20.3	135	2	T15610	hypothetical prote
10	98	20.2	137	2	T15609	hypothetical prote
11	96.5	19.9	869	1	JC4858	VLDL receptor prec
12	95.5	19.7	166	2	H89044	protein B0238.12 [
13	94.5	19.4	98	2	C89045	protein C1038.4 [
14	94.5	19.4	4560	2	T42737	gp330 protein prec
15	92.5	19.0	63	2	S07127	chymotrypsin/elast
16	92.5	19.0	63	2	S08572	chymotrypsin/elast
17	92.5	19.0	4753	1	A74737	LDL-receptor-relat
18	90.5	18.6	195	2	T28803	hypothetical prote
19	87.5	18.0	802	2	T24293	hypothetical prote
20	87.5	18.0	949	2	T24294	hypothetical prote
21	85.5	17.6	863	1	S51789	VLDL receptor prec
22	85	17.5	215	2	T30197	alpha tectorin - m
23	84.5	17.4	13288	2	T03099	muscin, submaxillar
24	83	17.1	209	2	T02394	hypothetical prote
25	82.5	17.0	873	1	A49729	VLDL receptor prec
26	82.5	17.0	873	1	QRREVD	trypsin inhibitor
27	82	16.9	62	2	S35098	hypothetical prote
28	82	16.8	267	2	T30007	hypothetical prote
29	81.5	16.8	1810	1	A32230	tenascin precursor

T27319  
 hypothetical protein Y69H2.3b - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #taxid 6233  
 C:Accession: T27319  
 R:McMurray, A.  
 submitted to the EMBL Data Library, August 1997  
 A:Reference number: Z20343  
 A:Accession: T27319  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-626 <WIL>  
 A:cross-references: EMBL:Z98877; PIDN:CAB54473.1; GSPDI  
 A:Experimental source: clone Y69H2  
 C:Genetics:  
 A:Gene: CESP:Y69H2.3b  
 A:Map position: 5  
 A:Intons: 183/1; 247/1; 312/1; 353/1; 383/1; 429/1; 4





A;Molecule type: mRNA  
A;Residues: 1-869 <OKA>  
A;Cross-references: GB:AB006906; NID:92366772; PIDN:BAA2145.1; PID:G2366773  
C;Comment: This receptor mediates incorporation of vitellogenin into oocytes.  
C;Superfamily: LDL receptor; EGF homology; LDL receptor ligand-binding repeat homology;  
C;Keywords: duplication; fatty acid metabolism; glycoprotein; receptor; transmembrane p  
F;1-26/Domain: signal sequence #status predicted <SIG>  
F;27-869/Product: VLDEL receptor #status predicted <SIG>  
F;27-793/Domain: extracellular #status predicted <EXT>  
F;32-66/Domain: LDL receptor ligand-binding repeat homology <LDL1>  
F;71-107/Domain: LDL receptor ligand-binding repeat homology <LDL2>  
F;112-148/Domain: LDL receptor ligand-binding repeat homology <LDL3>  
F;153-187/Domain: LDL receptor ligand-binding repeat homology <LDL4>  
F;192-228/Domain: LDL receptor ligand-binding repeat homology <LDL5>  
F;238-272/Domain: LDL receptor ligand-binding repeat homology <LDL6>  
F;277-311/Domain: LDL receptor ligand-binding repeat homology <LDL7>  
F;317-354/Domain: LDL receptor ligand-binding repeat homology <LDL8>  
F;359-393/Domain: EGF homology <EG1>  
F;399-433/Domain: EGF homology <EG2>  
F;440-479/Domain: LDL receptor YWTD-containing repeat homology <YWL1>  
F;480-525/Domain: LDL receptor YWTD-containing repeat homology <YWL2>  
F;526-568/Domain: LDL receptor YWTD-containing repeat homology <YWL3>  
F;569-612/Domain: LDL receptor YWTD-containing repeat homology <YWL4>  
F;613-655/Domain: LDL receptor YWTD-containing repeat homology <YWL5>  
F;656-698/Domain: LDL receptor YWTD-containing repeat homology <YWL6>  
F;707-749/Domain: EGF homology <EG3>  
F;794-815/Domain: transmembrane #status predicted <TM>  
F;816-869/Domain: intracellular #status predicted <CYT>  
F;830-834/Region: coated-pit mediated internalization signal  
F;150,201,777,786/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;359-370,366-379,381-393,399-409,405-418,420-433,707-720,716-735,737-749/Disulfide bond

Query Match 19.4%; Score 96.5; DB 1; Length 869;  
Best Local Similarity 35.4%; Pred. No. 0.16;  
Matches 28; Conservative 8; Mismatches 38; Indels 5; Gaps 4;

QY 2 ATMOCGENEKYDSCGSKEC-DKKCKYDGVVEEDDEPNVPCLVRVCHQD-CVCEGFGYRN 59  
DB 234 APORCSANEM--PCGSGECIHKWRCDGADCKDKSDEINCPSTCQDPQFCEDGNCIH 291

QY 60 KDDKCVSAEDCELDNMDFI 78  
DB 292 GSRQCDGVRDC-LDGTDEI 309

RESULT 12  
H89044  
protein B0238.12 [imported] - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C;Accession: H89044  
R;Anonymous, The C. elegans Sequencing Consortium.  
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog  
A;Reference number: A75000; MUID:99069613; PMID:9851916  
A;Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_ele  
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A;Accession: H89044  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-166 <STO>  
A;Cross-references: GB:chr\_V; PIDN:AAB65990.1; PID:G2315490; GSPDB:GN00023; CESP:B0238.1  
C;Genetics:  
A;Gene: B0238.12  
A;Map position: 5

Query Match 19.7%; Score 95.5; DB 2; Length 166;  
Best Local Similarity 31.4%; Pred. No. 0.049;  
Matches 22; Conservative 9; Mismatches 28; Indels 11; Gaps 4;

QY 1 KATMOGENEKYDSCGSKEC-DKKCKYDGVVEEDDEPNVPCLVRVCHQD-CVCEGFGYRN 60  
DB 108 QSPQRCGRNETRTGUS-SCBPSC-----TTPRPACTMQCLVNVQC-----CSEGFVRGP 157

QY 61 DDKCVSAEDC 70  
DB 158 SG-CVRQRDC 166

RESULT 13  
C89046  
protein C10G8.4 [imported] - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C;Accession: C89046  
R;Anonymous, The C. elegans Sequencing Consortium.  
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog  
A;Reference number: A75000; MUID:99069613; PMID:9851916  
A;Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_ele  
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A;Accession: C89046  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-98 <STO>  
A;Cross-references: GB:chr\_V; PIDN:AAB09171.1; PID:G1572829; GSPDB:GN00023; CESP:C10G8.4  
C;Genetics:  
A;Gene: C10G8.4  
A;Map position: 5

Query Match 19.4%; Score 94.5; DB 2; Length 98;  
Best Local Similarity 28.8%; Pred. No. 0.039;  
Matches 23; Conservative 13; Mismatches 17; Indels 27; Gaps 6;

QY 5 QCGENKDYDSCGSKEC-DKKCKYDGVVEEDDEPN-----VPCLVRVCHQD-CVCEGFGYRN 59  
DB 39 RCPNSNEPRSGCT-ACEPTC-----QNFNPQVCTLQCLVNVQC-----CSQGFVRG 83

QY 60 KDDKCVSAEDCELDNMDFI 79  
DB 84 PNG-CVPPQDC-----FVY 96

RESULT 14  
T42737  
gp330 protein precursor - rat  
N;Alternate names: megalin  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 04-Mar-2000  
C;Accession: T42737  
R;Saito, A.; Pietromonaco, S.; Loo, A.K.C.; Farquhar, M.G.  
Proc. Natl. Acad. Sci. U.S.A. 91, 9725-9729, 1994  
A;Title: Complete cloning and sequencing of rat gp330/megalin, a distinctive member of th  
A;Reference number: A58173; MUID:95024033; PMID:7937880  
A;Accession: T42737  
A;Status: preliminary; translated from GB/EMBL/DDJ  
A;Molecule type: mRNA  
A;Residues: 1-4660 <SAI>  
A;Cross-references: EMBL:L34049; NID:G561852; PID:G561853; PIDN:AAA51369.1  
A;Experimental source: strain Sprague-Dawley; kidney  
C;Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand-binding  
F;1-25/Domain: signal sequence #status predicted <SIG>  
F;26-4660/Product: gp330 protein #status predicted <MAT>

Query Match 19.4%; Score 94.5; DB 2; Length 4660;  
Best Local Similarity 30.0%; Pred. No. 0.96;  
Matches 30; Conservative 11; Mismatches 24; Indels 35; Gaps 8;

QY 6 CGENKDYDSC-----GSKEC---DKCKYDGVVEE---EDDEPNV-----PCLVRV--- 45  
DB 256 CESNQSHRCYCPREWACPGSGRCISDKVC--DGVPDCEGDENNVTSGRTCGMGVCSV 313

QY 46 -----CHQ-----DCVCEGFGYRNKDDK--CVSAEDCEL 72  
DB 314 LNCYQCHQTFPGGECFCPPGHIIINSNDSTCIDFDCCQI 353

RESULT 15

S07127  
 chymotrypsin/elastase inhibitor - common roundworm  
 C:Species: Ascaris lumbricoides (common roundworm)  
 C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 24-Jul-1997  
 C:Accession: S07127  
 R:Babin, D.R.; Pearsall, R.J.; Goos, S.M.  
 Arch. Biochem. Biophys. 232, 143-161, 1984  
 A:Title: The isoforms of chymotrypsin/elastase from Ascaris lumbricoides: the prima  
 A:Reference number: S07127; MUID:84255715; PMID:6564898  
 A:Accession: S07127  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-63 <BAR>  
 C:Superfamily: roundworm trypsin inhibitor

Query Match	19.0%	Score 92.5;	DB 2;	Length 63;
Best Local Similarity	32.8%	Pred. No. 0.042;		
Matches	22;	Conservative	5;	Mismatches 27; Indels 13; Gaps 4;
Qy	6	CGENEKYDSCGSKCEDKKKYDGVVEEDDEEPNVPCLVRYCHQDCVCE--EGFYRNKDDK	63	
Db	5	CGPNEVWTECTG--CEMKC---GPDE-----NTPCPLMCRPSCSCSPGRGWRRTNDGK	53	
Qy	64	CVSREDC	70	
Db	54	CIPASQC	60	

Search completed: September 24, 2004, 07:36:00  
 Job time : 53.2842 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 07:28:33 ; Search time 29.1789 Seconds  
(without alignments)  
149.899 Million cell updates/sec

Title: US-09-498-556C-59

Perfect score: 486

Sequence: 1 KATMQGGENEYDSCGSKC.....VSAEDCELDNDFIYPGTRN 84

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	117	24.1	77	1 CVP6_PIMHY	Q8T0W0 pimla hypo
2	110.5	22.7	56	1 AMCI_APIME	P56682 apis mellif
3	103	21.2	2812	1 ZAN_HUMAN	Q9V493 homo sapien
4	100.5	20.7	170	1 IMET_GALME	P82176 galleria me
5	100	20.6	71	1 ALL6_APIME	P83563 apis mellif
6	100	20.6	5376	1 ZAN_MOUSE	O88799 mus musculu
7	94.5	19.4	4660	1 LR2_RAT	P98158 rattus norv
8	92.5	19.0	63	1 ICB1_ASCSU	P07851 ascaris suu
9	92.5	19.0	4753	1 LRP_CAEEL	Q04833 caenorhabdi
10	87.5	18.0	2813	1 VWF_CANFA	Q28295 canis fami
11	87	17.9	115	1 A63E_DROME	Q48202 drosophila
12	86.5	17.8	4655	1 LR2_HUMAN	P98164 homo sapien
13	85.5	17.6	65	1 ICE2_ASCSU	P07852 ascaris suu
14	85.5	17.6	863	1 LDVR_CHICK	P98165 gallus gall
15	82.5	17.0	873	1 LDVR_HUMAN	P98156 homo sapien
16	82.5	17.0	873	1 LDVR_MOUSE	P98156 mus musculu
17	82.5	17.0	873	1 LDVR_RABIT	P35953 oryctolagus
18	82.5	17.0	873	1 LDVR_RAT	P98166 rattus norv
19	82.5	17.0	2282	1 ZAN_RABIT	P57999 oryctolagus
20	82	16.9	62	1 ITR1_ASCSU	P19398 ascaris suu
21	81.5	16.8	59	1 XCI_BOOMI	P83516 boophilus m
22	81.5	16.8	1808	1 META_CHICK	P10039 gallus gall
23	81	16.7	62	1 MT4_CANFA	Q9KUI5 canis fami
24	80	16.5	790	1 AD10_HUMAN	Q9UKF2 homo sapien
25	79.5	16.4	2813	1 VWF_HUMAN	P04275 homo sapien
26	79.5	16.4	2871	1 FBNI_HUMAN	P35555 homo sapien
27	79.5	16.4	2871	1 FBNI_PIG	Q9TV36 sus scrofa
28	78.5	16.2	77	1 ASP2_ANISI	O77417 anisakis si
29	78.5	16.2	798	1 ITB1_XENLA	P12606 xenopus lae
30	78.5	16.2	937	1 VWF_BOVIN	P80012 bos taurus
31	78.5	16.2	3375	1 UN52_CAEEL	O08561 caenorhabdi
32	78	16.2	289	1 TNR5_MOUSE	P27512 mus musculu
33	78	16.0	1107	1 YLK2_CAEEL	P41950 caenorhabdi

34	78	16.0	3133	1 HMCT_BOMMO	P98092 bombyx mori
35	77.5	15.9	474	1 VSM5_TRYBB	P26333 trypanosoma
36	77.5	15.9	2871	1 FBNI_BOVIN	P98133 bos taurus
37	77.5	15.9	2871	1 FBNI_MOUSE	Q61534 mus musculu
38	77	15.8	62	1 MT4_HUMAN	P47944 homo sapien
39	77	15.8	62	1 MT4_MOUSE	P47945 mus musculu
40	77	15.8	798	1 ITB1_MOUSE	P09055 mus musculu
41	76.5	15.7	66	1 ITR2_ASCSU	P01049 ascaris suu
42	76.5	15.7	1984	1 YL DROME	P98163 drosophila
43	76.5	15.7	2139	1 CRE_DROME	P10040 drosophila
44	76	15.6	799	1 ITB1_RAT	P49134 rattus norv
45	75.5	15.5	77	1 ASPI_ANISI	O77416 anisakis si

ALIGNMENTS

RESULT 1

CVP6\_PIMHY  
ID CVP6\_PIMHY STANDARD; PRT; 77 AA.  
AC Q8T0W0;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Cysteine-rich venom protein 6 precursor.  
GN CVP6.  
OS Pimla hypochondriaca (Parasitoid wasp).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Ichneumonidae;  
OC Ichneumonidae; Pimplinae; Pimpla.  
OX NCBI\_TaxID=135724;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Venom;  
RA Parkinson N.M.;  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).  
CC -!- SIMILARITY: Contains 1 TIL (Trypsin inhibitory-like) domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC EMBL; AJ438997; CAD27742.1; -  
CC InterPro; IPR002919; TIL\_Cysrich.  
CC Pfam; PF01826; TIL; 1.  
CC  
CC SIGNAL  
CC SIGNAL 1 21 Cysteine-rich venom protein 6.  
CC CHAIN 22 77 TIL.  
CC DOMAIN 24 76  
CC DISULFID 24 56 BY SIMILARITY.  
CC DISULFID 33 52 BY SIMILARITY.  
CC DISULFID 37 48 BY SIMILARITY.  
CC DISULFID 41 76 BY SIMILARITY.  
CC DISULFID 58 70 BY SIMILARITY.  
CC SEQUENCE 77 AA; 8184 MW; 6889C85F40D63DD6 CRC64;

Query Match 24.1%; Score 117; DB 1; Length 77;  
Best Local Similarity 35.4%; Pred. No. 4.7e-05;  
Matches 23; Conservative 8; Mismatches 22; Indels 12; Gaps 3;

Qy	6	CGENEKVDSCGSKCKKCKYGVGVEEDDEPNVPLVHVCHQDCVCEEGFYRNKDKCV	65
Db	24	CGPNRVYKSGTG-GCPETC-----ENPPDC-DRACHQGCFCGKLLQDIGNCI	71
Qy	66	SAEDC 70	
Db	72	SPDKC 76	

```

RESULT 2
AMCI_APIME STANDARD; PRT; 56 AA.
ID AMCI_APIME
AC P56682;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Chymotrypsin inhibitor (AMCI).
OS Apis mellifera (Honeybee).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC Apidae; Apis.
OX NCBI_TaxID=7460;
RN [1]
RP SEQUENCE AND STRUCTURE BY NMR.
RC TISSUE=Hemolymph; PubMed=10411628;
RX MEDLINE=9933935;
RA Bania J., Stachowiak D., Polanowski A.;
RT "Primary structure and properties of the cathepsin G/chymotrypsin
RT inhibitor from the larval hemolymph of Apis mellifera.";
RL Eur. J. Biochem. 262:680-687(1999).
CC -!- FUNCTION: Chymotrypsin and cathepsin G inhibitor.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Contains 1 TIL (Trypsin inhibitory-like) domain.
DR PDB; 1CCV; 12-MAR-99.
DR InterPro; IPR002919; TIL_Cysrich.
DR Pfam; PF01826; TIL; 1.
KW Serine protease inhibitor; 3D-structure.
FT DOMAIN 3 56
FT DISULFID 3 36
FT DISULFID 12 32
FT DISULFID 16 28
FT DISULFID 20 56
FT DISULFID 38 50
FT TURN 5 6
FT TURN 7 13
FT STRAND 15 16
FT STRAND 33 38
FT TURN 40 41
FT STRAND 43 45
FT TURN 46 48
FT STRAND 49 51
FT HELIX 53 56
SQ SEQUENCE 56 AA; 5973 MW; 092B2815AE62B7F CRC64;
Query Match 22.7%; Score 110.5; DB 1; Length 56;
Best Local Similarity 33.3%; Pred. NO. 0.00014;
Matches 22; Conservative 8; Mismatches 25; Indels 11; Gaps 2;
QY 5 QCGENKYSQSGKECDKKCKVDGVEEDDEBNVPCLVRVCHQVCCEEGFYRNKDDKC 64
DB 2 ECGFNEVFNTCGS-ACAPTC-----AOFKTRICTMQCRIGCCQCGEFLRNGGAC 50
QY 65 VSAEDC 70
DB 51 VLPENC 56
RESULT 3
ZAN_HUMAN STANDARD; PRT; 2812 AA.
ID ZAN_HUMAN
AC Q9Y493; Q00218; Q96L85; Q96L87; Q96L88; Q96L89; Q96L90;
AC Q9BXN9; Q9BZ83; Q9BZ84; Q9BZ85; Q9BZ86; Q9BZ87; Q9BZ88;
DT 16-OCT-2001 (Rel. 40, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Zonadhesin precursor.
ZAN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4; 5 AND 6).
RC TISSUE=Testis;
RA Cheung T.L., Wassler M.J., Cornwall G.A., Hardy D.M.;
RT "Multiple intra-species variants of human zonadhesin.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP PARTIAL SEQUENCE FROM N.A.
RX MEDLINE=99018118; PubMed=9799793;
RA Glockner G., Scherer S., Schattevoy R., Boright A.P., Weber J.,
RA Tsui L.-C., Rosenthal A.;
RT "Large-scale sequencing of two regions in human chromosome 7q22:
RT analysis of 650 kb of genomic sequence around the EPO and CUTL1 loci
RT reveals 17 genes.";
RL Genome Res. 8:1060-1073(1998).
RN [3]
RP SEQUENCE OF 1810-2812 FROM N.A. (ISOFORM 1).
RX MEDLINE=21138439; PubMed=11239002;
RA Wilson M.D., Riemer C., Martindale D.W., Schnupf P., Boright A.P.,
RA Cheung T.L., Hardy D.M., Schwartz S., Scherer S.W., Tsui L.-C.,
RA Miller W., Koop B.F.;
RT "Comparative analysis of the gene-dense ACHE/TFR2 region on human
RT chromosome 7q22 with the orthologous region on mouse chromosome 5.";
RL Nucleic Acids Res. 29:1352-1365(2001).
RN [4]
RP SEQUENCE OF 2375-2683 FROM N.A. (ISOFORM 7).
RC TISSUE=Testis;
RX MEDLINE=97271566; PubMed=9126492;
RA Gao Z., Hartumi T., Garbers D.L.;
RT "Chromosome localization of the mouse zonadhesin gene and the human
RT zonadhesin gene (ZAN)";
RL Genomics 41:119-122(1997).
CC -!- FUNCTION: Binds in a species-specific manner to the zona pellucida
CC of the egg. May be involved in gamete recognition and/or
CC signaling.
CC -!- SUBUNIT: Probably forms covalent oligomers.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein, exclusively on the
CC apical region of the sperm head (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=7;
CC Name=3;
CC IsoId=Q9Y493-1; Sequence=Displayed;
CC Name=1;
CC IsoId=Q9Y493-2; Sequence=VSP_001430, VSP_001431;
CC Name=2;
CC IsoId=Q9Y493-3; Sequence=VSP_001428, VSP_001429;
CC Name=4;
CC IsoId=Q9Y493-4; Sequence=VSP_001424, VSP_001425;
CC Name=5;
CC IsoId=Q9Y493-5; Sequence=VSP_001420, VSP_001421;
CC Name=6;
CC IsoId=Q9Y493-6; Sequence=VSP_001422, VSP_001423;
CC Name=7;
CC IsoId=Q9Y493-7; Sequence=VSP_001426, VSP_001427;
CC -!- TISSUE SPECIFICITY: In testis, primarily in haploid spermatids.
CC -!- DOMAIN: The ZAN domains probably mediate sperm adhesion to the
CC zona pellucida.
CC -!- DOMAIN: During sperm migration through the reproductive tracts,
CC the mucin-like domain might inhibit inappropriate trapping of
CC spermatozoa or promoting adhesion to the oviductal isthmus.
CC -!- DOMAIN: The VWFD domain 2 may mediate covalent oligomerization (by
CC similarity to human intestinal mucin MUC2).
CC -!- SIMILARITY: Contains 3 ZAN domains.
CC -!- SIMILARITY: Contains 4 VWFD domains.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- CAUTION: Ref.2 sequence differs from that shown due to
CC transposition of a number of exons.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial

```





RX MEDLINE-97271566; PubMed-9126492;  
 RA Gao Z., Harumi T., Garbers D.L.;  
 RT "Chromosome localization of the mouse zonadhesin gene and the human  
 RT zonadhesin gene (ZAN).";  
 RL Genomics 41:119-122(1997).  
 CC -!- FUNCTION: Binds in a species-specific manner to the zona pellucida  
 CC of the egg. May be involved in gamete recognition and/or  
 CC signaling.  
 CC -!- SUBUNIT: Probably forms covalent oligomers.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein, exclusively on the  
 CC apical region of the sperm head.  
 CC -!- TISSUE SPECIFICITY: In testis, primarily in haploid spermatids.  
 CC -!- DOMAIN: The MAM domains probably mediates sperm adhesion to the  
 CC zona pellucida.  
 CC -!- DOMAIN: During sperm migration through the reproductive tracts,  
 CC the mucin-like domain might inhibit inappropriate trapping of  
 CC spermatozoa or promoting adhesion to the oviductal isthmus.  
 CC -!- DOMAIN: The VWFD domain 2 may mediate covalent oligomerization (by  
 CC similarity to human intestinal mucin MUC2).  
 CC -!- SIMILARITY: Contains 3 MAM domains.  
 CC -!- SIMILARITY: Contains 25 VWFD domains.  
 CC -!- SIMILARITY: Contains 1 EGF-like domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U97068; AAC26680.1; -;  
 CC EMBL; U83190; AAC53125.1; -;  
 CC PIR; T42215; T42215  
 CC MGD; MGI:106556; Zan.  
 CC InterPro; IPR006209; EGF like.  
 CC InterPro; IPR003645; FOLN.  
 CC InterPro; IPR006210; IEGF.  
 CC InterPro; IPR000998; MAM domain.  
 CC InterPro; IPR002919; TIL Cysrich.  
 CC InterPro; IPR003328; TILA Cysrich.  
 CC InterPro; IPR001007; VWF C.  
 CC InterPro; IPR001846; VWF\_D.  
 CC Pfam; PF00629; MAM; 3.  
 CC Pfam; PF01826; TIL; 25.  
 CC Pfam; PF02345; TILA; 25.  
 CC Pfam; PF00094; vwd; 4.  
 CC SMART; SM00181; EGF; 2.  
 CC SMART; SM00274; FOLN; 11.  
 CC SMART; SM00137; MAM; 2.  
 CC SMART; SM00214; VWC; 17.  
 CC SMART; SM00216; VWD; 4.  
 CC PROSITE; PS00022; EGF 1; 1.  
 CC PROSITE; PS01186; EGF 2; 18.  
 CC PROSITE; PS00026; EGF 3; 1.  
 CC PROSITE; PS00740; MAM\_1; FALSE\_NEG.  
 CC PROSITE; PS00600; MAM\_2; 3.  
 KW Signal; Glycoprotein; Transmembrane; EGF-like domain; Cell adhesion;  
 KW Repeat.  
 FT SIGNAL 1 17 POTENTIAL.  
 FT CHAIN 18 5376 ZONADHESIN.  
 FT DOMAIN 18 5310 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 5311 5337 POTENTIAL.  
 FT DOMAIN 5338 5376 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 45 210 MAM 1.  
 FT DOMAIN 215 374 MAM 2.  
 FT DOMAIN 377 542 MAM 3.  
 FT DOMAIN 547 1170 80 X HEPTAPEPTIDE REPEATS (APPROXIMATE)  
 FT (MUCIN-LIKE DOMAIN).  
 FT DOMAIN 1171 1280 VWFD 1 (PARTIAL).  
 FT DOMAIN 1281 1669 VWFD 2.  
 FT DOMAIN 1670 2056 VWFD 3.  
 FT DOMAIN 2057 2459 VWFD 4.

FT DOMAIN	2460	2579	VWFD 5 (PARTIAL).
FT DOMAIN	2580	2699	VWFD 6 (PARTIAL).
FT DOMAIN	2819	2819	VWFD 7 (PARTIAL).
FT DOMAIN	2820	2939	VWFD 8 (PARTIAL).
FT DOMAIN	2940	3059	VWFD 9 (PARTIAL).
FT DOMAIN	3060	3179	VWFD 10 (PARTIAL).
FT DOMAIN	3180	3299	VWFD 11 (PARTIAL).
FT DOMAIN	3300	3416	VWFD 12 (PARTIAL).
FT DOMAIN	3417	3536	VWFD 13 (PARTIAL).
FT DOMAIN	3537	3656	VWFD 14 (PARTIAL).
FT DOMAIN	3657	3776	VWFD 15 (PARTIAL).
FT DOMAIN	3777	3892	VWFD 16 (PARTIAL).
FT DOMAIN	3893	4928	VWFD 17 (PARTIAL).
FT DOMAIN	4029	4148	VWFD 18 (PARTIAL).
FT DOMAIN	4149	4263	VWFD 19 (PARTIAL).
FT DOMAIN	4264	4283	VWFD 20 (PARTIAL).
FT DOMAIN	4384	4503	VWFD 21 (PARTIAL).
FT DOMAIN	4504	4623	VWFD 22 (PARTIAL).
FT DOMAIN	4624	4743	VWFD 23 (PARTIAL).
FT DOMAIN	4744	4863	VWFD 24 (PARTIAL).
FT DOMAIN	4864	5261	VWFD 25.
FT DOMAIN	5259	5295	EGF-LIKE.
FT DISULFID	5263	5274	BY SIMILARITY.
FT DISULFID	5268	5283	BY SIMILARITY.
FT DISULFID	5285	5294	BY SIMILARITY.
FT CARBOHYD	339	339	N-LINKED (GLCNAC. .)
FT CARBOHYD	499	499	N-LINKED (GLCNAC. .)
FT CARBOHYD	1216	1216	N-LINKED (GLCNAC. .)
FT CARBOHYD	1239	1239	N-LINKED (GLCNAC. .)
FT CARBOHYD	1314	1314	N-LINKED (GLCNAC. .)
FT CARBOHYD	1814	1814	N-LINKED (GLCNAC. .)
FT CARBOHYD	1908	1908	N-LINKED (GLCNAC. .)
FT CARBOHYD	1933	1933	N-LINKED (GLCNAC. .)
FT CARBOHYD	2028	2028	N-LINKED (GLCNAC. .)
FT CARBOHYD	2111	2111	N-LINKED (GLCNAC. .)
FT CARBOHYD	2142	2142	N-LINKED (GLCNAC. .)
FT CARBOHYD	2332	2332	N-LINKED (GLCNAC. .)
FT CARBOHYD	2533	2533	N-LINKED (GLCNAC. .)
FT CARBOHYD	2575	2575	N-LINKED (GLCNAC. .)
FT CARBOHYD	2692	2692	N-LINKED (GLCNAC. .)
FT CARBOHYD	2812	2812	N-LINKED (GLCNAC. .)
FT CARBOHYD	3052	3052	N-LINKED (GLCNAC. .)
FT CARBOHYD	3065	3065	N-LINKED (GLCNAC. .)
FT CARBOHYD	3144	3144	N-LINKED (GLCNAC. .)
FT CARBOHYD	3172	3172	N-LINKED (GLCNAC. .)
FT CARBOHYD	3288	3288	N-LINKED (GLCNAC. .)
FT CARBOHYD	3292	3292	N-LINKED (GLCNAC. .)
FT CARBOHYD	3782	3782	N-LINKED (GLCNAC. .)
FT CARBOHYD	4005	4005	N-LINKED (GLCNAC. .)
FT CARBOHYD	4136	4136	N-LINKED (GLCNAC. .)
FT CARBOHYD	4243	4243	N-LINKED (GLCNAC. .)
FT CARBOHYD	4254	4254	N-LINKED (GLCNAC. .)
FT CARBOHYD	4335	4335	N-LINKED (GLCNAC. .)
FT CARBOHYD	4376	4376	N-LINKED (GLCNAC. .)
FT CARBOHYD	4586	4586	N-LINKED (GLCNAC. .)
FT CARBOHYD	5136	5136	N-LINKED (GLCNAC. .)
FT CARBOHYD	5252	5252	N-LINKED (GLCNAC. .)
SQ SEQUENCE	5376 AA;	579908 MW;	0544D577DF2A2620 CRC64;

Query Match 20.6%; Score 100; DB 1; Length 5376;  
 Best Local Similarity 33.8%; Pred No. 0.15;  
 Matches 25; Conservative 10; Mismatches 27; Indels 12; Gaps 5;  
 Qy 1 KATMQCGENKDYSCGSKCKCKYDGVVEE---DDEENPVLVRVCHQDCVCEGFY 57  
 Db 3412 KITLQCPAHTQYTSCLPSC---LDPEGLCKDIPKVT---STCKEGVCQSGYV 3463  
 Qy 58 RNKDDKCVSAEDCE 71  
 Db 3464 LN-SDRCVLRACD 3476  
 RESULT 7

LRP2 RAT  
ID LRP2 RAT STANDARD; PRT; 4660 AA.  
AC P8158;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Low-density lipoprotein receptor-related protein 2 precursor (Megalyn)  
DE (Glycoprotein 330) (gp330).  
GN LRP2.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=Sprague-Dawley; TISSUE=Kidney;  
RX MEDLINE=95024033; PubMed=7937880;  
RA Saito A., Pietromonaco S., Loo A.K.C., Farquhar M.G.;  
RT "Complete cloning and sequencing of rat gp330/megalin," a  
RT distinctive member of the low density lipoprotein receptor gene  
RT family.";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:9725-9729(1994).  
RN [2]  
RN FUNCTION  
RX MEDLINE=95386696; PubMed=7544804;  
RA Moestrup S.K., Cui S., Vorum H., Bregengaard C., Bjorn S.E.,  
RA Norris K., Gliemann J., Christensen E.I.;  
RT "Evidence that epithelial glycoprotein 330/megalin mediates uptake of  
RT polybasic drugs.";  
RL J. Clin. Invest. 96:1404-1413(1995).  
RN [3]  
RN TISSUE SPECIFICITY  
RX MEDLINE=94172242; PubMed=7510321;  
RA Zheng G., Bachinsky R.T., Stamenkovic I., Strickland D.K., Brown D.,  
RA Andres G., McCluskey R.T.;  
RT "Organ distribution in rats of two members of the low-density  
RT lipoprotein receptor gene family, gp330 and LRP/alpha 2MR, and the  
RT receptor-associated protein (RAP).";  
RL J. Histochem. Cytochem. 42:531-542(1994).  
CC -1- FUNCTION: Binds plasminogen, extracellular matrix components,  
CC plasminogen activator-plasminogen activator inhibitor type I  
CC complex, apolipoprotein B-enriched beta-VLDL, lipoprotein lipase,  
CC lactoferrin, clusterin and calcium.  
CC -1- FUNCTION: Receptor-mediated uptake of polybasic drugs such as  
CC aprotinin, aminoglycosides and polymyxin B.  
CC -1- SUBUNIT: Forms a multimeric complex together with a receptor-  
CC associated protein (RAP). Binds to ankryrin-repeat family A protein  
CC 2 (ANKRA2) (By similarity).  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Expressed in  
CC clathrin-coated pits; a soluble form is possibly derived by  
CC cleavage at the cell surface.  
CC -1- TISSUE SPECIFICITY: Epithelial cells of kidney glomerulus and  
CC proximal tubule, lung, epididymis, yolk sac, among others.  
CC -1- SIMILARITY: Contains 36 LDL-receptor class A domains.  
CC -1- SIMILARITY: Contains 37 LDL-receptor class B domains.  
CC -1- SIMILARITY: Contains 17 EGF-like domains.  
CC -----  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L34049; AAA51369.1;  
DR PIR; T42737; T42737.  
DR HSP; Q07954; ICR8.  
DR GlycosuitedB; P98158;  
DR InterPro; IPR000152; Asx\_hydroxyl\_S.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR006209; EGF\_Like.  
DR InterPro; IPR002172; LDL\_receptor\_A.

InterPro; IPR000033; Ldl\_receptor\_rep.  
DR Pfam; PF00008; EGF; 9.  
DR Pfam; PF00057; ldl\_recept\_a; 36.  
DR Pfam; PF00058; ldl\_recept\_b; 33.  
DR PRINTS; PR00261; LDLRECEPTOR.  
DR SMART; SMO0179; EGF\_CA; 3.  
DR SMART; SMO0192; LDLA; 36.  
DR SMART; SMO0135; LY; 34.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 4.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS01186; EGF\_2; 8.  
DR PROSITE; PS00026; EGF\_3; 8.  
DR PROSITE; PS01187; EGF\_CA; 3.  
DR PROSITE; PS01209; LDLA\_1; 31.  
DR PROSITE; PS00068; LDLA\_2; 36.  
KW Glycoprotein; Repeat; Endocytosis; Coated pits; Transmembrane;  
KW Receptor; EGF-like domain; SH3-binding; Signal.  
FT SIGNAL 1 25  
FT CHAIN 26 4660  
FT DOMAIN 26 4425  
FT TRANSMEM 4426 4446  
FT DOMAIN 4447 4660  
FT DOMAIN 26 64  
FT DOMAIN 65 105  
FT DOMAIN 106 144  
FT DOMAIN 145 181  
FT DOMAIN 181 219  
FT DOMAIN 220 258  
FT DOMAIN 263 307  
FT DOMAIN 308 346  
FT DOMAIN 347 385  
FT DOMAIN 435 476  
FT DOMAIN 478 519  
FT DOMAIN 521 566  
FT DOMAIN 568 611  
FT DOMAIN 612 652  
FT DOMAIN 658 703  
FT DOMAIN 752 793  
FT DOMAIN 795 835  
FT DOMAIN 837 879  
FT DOMAIN 881 923  
FT DOMAIN 969 1013  
FT DOMAIN 1023 1061  
FT DOMAIN 1064 1103  
FT DOMAIN 1108 1146  
FT DOMAIN 1148 1186  
FT DOMAIN 1186 1225  
FT DOMAIN 1229 1269  
FT DOMAIN 1270 1308  
FT DOMAIN 1311 1351  
FT DOMAIN 1350 1390  
FT DOMAIN 1391 1430  
FT DOMAIN 1479 1520  
FT DOMAIN 1522 1563  
FT DOMAIN 1567 1609  
FT DOMAIN 1611 1654  
FT DOMAIN 1656 1696  
FT DOMAIN 1701 1742  
FT DOMAIN 1791 1832  
FT DOMAIN 1834 1882  
FT DOMAIN 1884 1930  
FT DOMAIN 1932 1972  
FT DOMAIN 1973 2013  
FT DOMAIN 2019 2060  
FT DOMAIN 2108 2156  
FT DOMAIN 2158 2201  
FT DOMAIN 2203 2245  
FT DOMAIN 2247 2289  
FT DOMAIN 2291 2332  
FT DOMAIN 2343 2384  
FT DOMAIN 2432 2477  
FT DOMAIN 2479 2518  
FT DOMAIN 2519 2560  
FT DOMAIN 2561 2602  
FT DOMAIN 2603 2644  
FT DOMAIN 2645 2686  
FT DOMAIN 2687 2728  
FT DOMAIN 2729 2770  
FT DOMAIN 2771 2812  
FT DOMAIN 2813 2854  
FT DOMAIN 2855 2896  
FT DOMAIN 2897 2938  
FT DOMAIN 2939 2980  
FT DOMAIN 2981 3022  
FT DOMAIN 3023 3064  
FT DOMAIN 3065 3106  
FT DOMAIN 3107 3148  
FT DOMAIN 3149 3190  
FT DOMAIN 3191 3232  
FT DOMAIN 3233 3274  
FT DOMAIN 3275 3316  
FT DOMAIN 3317 3358  
FT DOMAIN 3359 3400  
FT DOMAIN 3401 3442  
FT DOMAIN 3443 3484  
FT DOMAIN 3485 3526  
FT DOMAIN 3527 3568  
FT DOMAIN 3569 3610  
FT DOMAIN 3611 3652  
FT DOMAIN 3653 3694  
FT DOMAIN 3695 3736  
FT DOMAIN 3737 3778  
FT DOMAIN 3779 3820  
FT DOMAIN 3821 3862  
FT DOMAIN 3863 3904  
FT DOMAIN 3905 3946  
FT DOMAIN 3947 3988  
FT DOMAIN 3989 4030  
FT DOMAIN 4031 4072  
FT DOMAIN 4073 4114  
FT DOMAIN 4115 4156  
FT DOMAIN 4157 4198  
FT DOMAIN 4199 4240  
FT DOMAIN 4241 4282  
FT DOMAIN 4283 4324  
FT DOMAIN 4325 4366  
FT DOMAIN 4367 4408  
FT DOMAIN 4409 4450  
FT DOMAIN 4451 4492  
FT DOMAIN 4493 4534  
FT DOMAIN 4535 4576  
FT DOMAIN 4577 4618  
FT DOMAIN 4619 4660  
FT DOMAIN 4661 4702  
FT DOMAIN 4703 4744  
FT DOMAIN 4745 4786  
FT DOMAIN 4787 4828  
FT DOMAIN 4829 4870  
FT DOMAIN 4871 4912  
FT DOMAIN 4913 4954  
FT DOMAIN 4955 4996  
FT DOMAIN 4997 5038  
FT DOMAIN 5039 5080  
FT DOMAIN 5081 5122  
FT DOMAIN 5123 5164  
FT DOMAIN 5165 5206  
FT DOMAIN 5207 5248  
FT DOMAIN 5249 5290  
FT DOMAIN 5291 5332  
FT DOMAIN 5333 5374  
FT DOMAIN 5375 5416  
FT DOMAIN 5417 5458  
FT DOMAIN 5459 5500  
FT DOMAIN 5501 5542  
FT DOMAIN 5543 5584  
FT DOMAIN 5585 5626  
FT DOMAIN 5627 5668  
FT DOMAIN 5669 5710  
FT DOMAIN 5711 5752  
FT DOMAIN 5753 5794  
FT DOMAIN 5795 5836  
FT DOMAIN 5837 5878  
FT DOMAIN 5879 5920  
FT DOMAIN 5921 5962  
FT DOMAIN 5963 6004  
FT DOMAIN 6005 6046  
FT DOMAIN 6047 6088  
FT DOMAIN 6089 6130  
FT DOMAIN 6131 6172  
FT DOMAIN 6173 6214  
FT DOMAIN 6215 6256  
FT DOMAIN 6257 6298  
FT DOMAIN 6299 6340  
FT DOMAIN 6341 6382  
FT DOMAIN 6383 6424  
FT DOMAIN 6425 6466  
FT DOMAIN 6467 6508  
FT DOMAIN 6509 6550  
FT DOMAIN 6551 6592  
FT DOMAIN 6593 6634  
FT DOMAIN 6635 6676  
FT DOMAIN 6677 6718  
FT DOMAIN 6719 6760  
FT DOMAIN 6761 6802  
FT DOMAIN 6803 6844  
FT DOMAIN 6845 6886  
FT DOMAIN 6887 6928  
FT DOMAIN 6929 6970  
FT DOMAIN 6971 7012  
FT DOMAIN 7013 7054  
FT DOMAIN 7055 7096  
FT DOMAIN 7097 7138  
FT DOMAIN 7139 7180  
FT DOMAIN 7181 7222  
FT DOMAIN 7223 7264  
FT DOMAIN 7265 7306  
FT DOMAIN 7307 7348  
FT DOMAIN 7349 7390  
FT DOMAIN 7391 7432  
FT DOMAIN 7433 7474  
FT DOMAIN 7475 7516  
FT DOMAIN 7517 7558  
FT DOMAIN 7559 7600  
FT DOMAIN 7601 7642  
FT DOMAIN 7643 7684  
FT DOMAIN 7685 7725  
FT DOMAIN 7726 7767  
FT DOMAIN 7768 7809  
FT DOMAIN 7810 7851  
FT DOMAIN 7852 7893  
FT DOMAIN 7894 7935  
FT DOMAIN 7936 7977  
FT DOMAIN 7978 8019  
FT DOMAIN 8020 8061  
FT DOMAIN 8062 8103  
FT DOMAIN 8104 8145  
FT DOMAIN 8146 8187  
FT DOMAIN 8188 8229  
FT DOMAIN 8230 8271  
FT DOMAIN 8272 8313  
FT DOMAIN 8314 8355  
FT DOMAIN 8356 8397  
FT DOMAIN 8398 8439  
FT DOMAIN 8440 8481  
FT DOMAIN 8482 8523  
FT DOMAIN 8524 8565  
FT DOMAIN 8566 8607  
FT DOMAIN 8608 8649  
FT DOMAIN 8650 8691  
FT DOMAIN 8692 8733  
FT DOMAIN 8734 8775  
FT DOMAIN 8776 8817  
FT DOMAIN 8818 8859  
FT DOMAIN 8860 8901  
FT DOMAIN 8902 8943  
FT DOMAIN 8944 8985  
FT DOMAIN 8986 9027  
FT DOMAIN 9028 9069  
FT DOMAIN 9070 9111  
FT DOMAIN 9112 9153  
FT DOMAIN 9154 9195  
FT DOMAIN 9196 9237  
FT DOMAIN 9238 9279  
FT DOMAIN 9280 9321  
FT DOMAIN 9322 9363  
FT DOMAIN 9364 9405  
FT DOMAIN 9406 9447  
FT DOMAIN 9448 9489  
FT DOMAIN 9490 9531  
FT DOMAIN 9532 9573  
FT DOMAIN 9574 9615  
FT DOMAIN 9616 9657  
FT DOMAIN 9658 9699  
FT DOMAIN 9700 9741  
FT DOMAIN 9742 9783  
FT DOMAIN 9784 9825  
FT DOMAIN 9826 9867  
FT DOMAIN 9868 9909  
FT DOMAIN 9910 9951  
FT DOMAIN 9952 9993  
FT DOMAIN 9994 10035  
FT DOMAIN 10036 10077  
FT DOMAIN 10078 10119  
FT DOMAIN 10120 10161  
FT DOMAIN 10162 10203  
FT DOMAIN 10204 10245  
FT DOMAIN 10246 10287  
FT DOMAIN 10288 10329  
FT DOMAIN 10330 10371  
FT DOMAIN 10372 10413  
FT DOMAIN 10414 10455  
FT DOMAIN 10456 10497  
FT DOMAIN 10498 10539  
FT DOMAIN 10540 10581  
FT DOMAIN 10582 10623  
FT DOMAIN 10624 10665  
FT DOMAIN 10666 10707  
FT DOMAIN 10708 10749  
FT DOMAIN 10750 10791  
FT DOMAIN 10792 10833  
FT DOMAIN 10834 10875  
FT DOMAIN 10876 10917  
FT DOMAIN 10918 10959  
FT DOMAIN 10960 11001  
FT DOMAIN 11002 11043  
FT DOMAIN 11044 11085  
FT DOMAIN 11086 11127  
FT DOMAIN 11128 11169  
FT DOMAIN 11170 11211  
FT DOMAIN 11212 11253  
FT DOMAIN 11254 11295  
FT DOMAIN 11296 11337  
FT DOMAIN 11338 11379  
FT DOMAIN 11380 11421  
FT DOMAIN 11422 11463  
FT DOMAIN 11464 11505  
FT DOMAIN 11506 11547  
FT DOMAIN 11548 11589  
FT DOMAIN 11590 11631  
FT DOMAIN 11632 11673  
FT DOMAIN 11674 11715  
FT DOMAIN 11716 11757  
FT DOMAIN 11758 11799  
FT DOMAIN 11800 11841  
FT DOMAIN 11842 11883  
FT DOMAIN 11884 11925  
FT DOMAIN 11926 11967  
FT DOMAIN 11968 12009  
FT DOMAIN 12010 12051  
FT DOMAIN 12052 12093  
FT DOMAIN 12094 12135  
FT DOMAIN 12136 12177  
FT DOMAIN 12178 12219  
FT DOMAIN 12220 12261  
FT DOMAIN 12262 12303  
FT DOMAIN 12304 12345  
FT DOMAIN 12346 12387  
FT DOMAIN 12388 12429  
FT DOMAIN 12430 12471  
FT DOMAIN 12472 12513  
FT DOMAIN 12514 12555  
FT DOMAIN 12556 12597  
FT DOMAIN 12598 12639  
FT DOMAIN 12640 12681  
FT DOMAIN 12682 12723  
FT DOMAIN 12724 12765  
FT DOMAIN 12766 12807  
FT DOMAIN 12808 12849  
FT DOMAIN 12850 12891  
FT DOMAIN 12892 12933  
FT DOMAIN 12934 12975  
FT DOMAIN 12976 13017  
FT DOMAIN 13018 13059  
FT DOMAIN 13060 13101  
FT DOMAIN 13102 13143  
FT DOMAIN 13144 13185  
FT DOMAIN 13186 13227  
FT DOMAIN 13228 13269  
FT DOMAIN 13270 13311  
FT DOMAIN 13312 13353  
FT DOMAIN 13354 13395  
FT DOMAIN 13396 13437  
FT DOMAIN 13438 13479  
FT DOMAIN 13480 13521  
FT DOMAIN 13522 13563  
FT DOMAIN 13564 13605  
FT DOMAIN 13606 13647  
FT DOMAIN 13648 13689  
FT DOMAIN 13690 13731  
FT DOMAIN 13732 13773  
FT DOMAIN 13774 13815  
FT DOMAIN 13816 13857  
FT DOMAIN 13858 13899  
FT DOMAIN 13900 13941  
FT DOMAIN 13942 13983  
FT DOMAIN 13984 14025  
FT DOMAIN 14026 14067  
FT DOMAIN 14068 14109  
FT DOMAIN 14110 14151  
FT DOMAIN 14152 14193  
FT DOMAIN 14194 14235  
FT DOMAIN 14236 14277  
FT DOMAIN 14278 14319  
FT DOMAIN 14320 14361  
FT DOMAIN 14362 14403  
FT DOMAIN 14404 14445  
FT DOMAIN 14446 14487  
FT DOMAIN 14488 14529  
FT DOMAIN 14530 14571  
FT DOMAIN 14572 14613  
FT DOMAIN 14614 14655  
FT DOMAIN 14656 14697  
FT DOMAIN 14698 14739  
FT DOMAIN 14740 14781  
FT DOMAIN 14782 14823  
FT DOMAIN 14824 14865  
FT DOMAIN 14866 14907  
FT DOMAIN 14908 14949  
FT DOMAIN 14950 14991  
FT DOMAIN 14992 15033  
FT DOMAIN 15034 15075  
FT DOMAIN 15076 15117  
FT DOMAIN 15118 15159  
FT DOMAIN 15160 15201  
FT DOMAIN 15202 15243  
FT DOMAIN 15244 15285  
FT DOMAIN 15286 15327  
FT DOMAIN 15328 15369  
FT DOMAIN 15370 15411  
FT DOMAIN 15412 15453  
FT DOMAIN 15454 15495  
FT DOMAIN 15496 15537  
FT DOMAIN 15538 15579  
FT DOMAIN 15580 15621  
FT DOMAIN 15622 15663  
FT DOMAIN 15664 15705  
FT DOMAIN 15706 15747  
FT DOMAIN 15748 15789  
FT DOMAIN 15790 15831  
FT DOMAIN 15832 15873  
FT DOMAIN 15874 15915  
FT DOMAIN 15916 15957  
FT DOMAIN 15958 15999  
FT DOMAIN 16000 16041  
FT DOMAIN 16042 16083  
FT DOMAIN 16084 16125  
FT DOMAIN 16126 16167  
FT DOMAIN 16168 16209  
FT DOMAIN 16210 16251  
FT DOMAIN 16252 16293  
FT DOMAIN 16294 16335  
FT DOMAIN 16336 16377  
FT DOMAIN 16378 16419  
FT DOMAIN 16420 16461  
FT DOMAIN 16462 16503  
FT DOMAIN 16504 16545  
FT DOMAIN 16546 16587  
FT DOMAIN 16588 16629  
FT DOMAIN 16630 16671  
FT DOMAIN 16672 16713  
FT DOMAIN 16714 16755  
FT DOMAIN 16756 16797  
FT DOMAIN 16798 16839  
FT DOMAIN 16840 16881  
FT DOMAIN 16882 16923  
FT DOMAIN 16924 16965  
FT DOMAIN 16966 17007  
FT DOMAIN 17008 17049  
FT DOMAIN 17050 17091  
FT DOMAIN 17092 17133  
FT DOMAIN 17134 17175  
FT DOMAIN 17176 17217  
FT DOMAIN 17218 17259  
FT DOMAIN 17260 17301  
FT DOMAIN 17302 17343  
FT DOMAIN 17344 17385  
FT DOMAIN 17386 17427  
FT DOMAIN 17428 17469  
FT DOMAIN 17470 17511  
FT DOMAIN 17512 17553  
FT DOMAIN 17554 17595  
FT DOMAIN 17596 17637  
FT DOMAIN 17638 17679  
FT DOMAIN 17680 17721  
FT DOMAIN 17722 17763  
FT DOMAIN 17764 17805  
FT DOMAIN 17806 17847  
FT DOMAIN 17848 17889  
FT DOMAIN 17890 17931  
FT DOMAIN 17932 17973  
FT DOMAIN 17974 18015  
FT DOMAIN 18016 18057  
FT DOMAIN 18058 18099  
FT DOMAIN 18100 18141  
FT DOMAIN 18142 18183  
FT DOMAIN 18184 18225  
FT DOMAIN 18226 18267  
FT DOMAIN 18268 18309  
FT DOMAIN 18310 18351  
FT DOMAIN 18352 18393  
FT DOMAIN 18394 18435  
FT DOMAIN 18436 18477  
FT DOMAIN 18478 18519  
FT DOMAIN 18520 18561  
FT DOMAIN 18562 18603  
FT DOMAIN 18604 18645  
FT DOMAIN 18646 18687  
FT DOMAIN 18688 18729  
FT DOMAIN 18730 18771  
FT DOMAIN 18772 18813  
FT DOMAIN 18814 18855  
FT DOMAIN 18856 18897  
FT DOMAIN 18898 18939  
FT DOMAIN 18940 18981  
FT DOMAIN 18982 19023  
FT DOMAIN 19024 19065  
FT DOMAIN 19066 19107  
FT DOMAIN 19108 19149  
FT DOMAIN 19150 19191  
FT DOMAIN 19192 19233  
FT DOMAIN 19234 19275  
FT DOMAIN 19276 19317  
FT DOMAIN 19318 19359  
FT DOMAIN 19360 19401  
FT DOMAIN 19402 19443  
FT DOMAIN 19444 19485  
FT DOMAIN 19486 19527  
FT DOMAIN 19528 19569  
FT DOMAIN 19570 19611  
FT DOMAIN 19612 19653  
FT DOMAIN 19654 19695  
FT DOMAIN 19696 19737  
FT DOMAIN 19738 19779  
FT DOMAIN 19780 19821  
FT DOMAIN 19822 19863  
FT DOMAIN 19864 19905  
FT DOMAIN 19906 19947  
FT DOMAIN 19948 19989  
FT DOMAIN 19990 20031  
FT DOMAIN 20032 20073  
FT DOMAIN 20074 20115  
FT DOMAIN 20116 20157  
FT DOMAIN 20158 20199  
FT DOMAIN 20200 20241  
FT DOMAIN 20242 20283  
FT DOMAIN 20284 20325  
FT DOMAIN 20326 20367  
FT DOMAIN 20368 20409  
FT DOMAIN 20410 20451  
FT DOMAIN 20452 20493  
FT DOMAIN 20494 20535  
FT DOMAIN 20536 20577  
FT DOMAIN 20578 20619  
FT DOMAIN 20620 20661  
FT DOMAIN 20662 20703  
FT DOMAIN 20704 20745  
FT DOMAIN 20746 20787  
FT DOMAIN 20788 20829  
FT DOMAIN 20830 20871  
FT DOMAIN 20872 20913  
FT DOMAIN 20914 20955  
FT DOMAIN 20956 20997  
FT DOMAIN 21000 21041  
FT DOMAIN 21042 21083  
FT DOMAIN 21084 21125  
FT DOMAIN 21126 21167  
FT DOMAIN 21168 21209  
FT DOMAIN 21210 21251  
FT DOMAIN 21252 21293  
FT DOMAIN 21294 21335  
FT DOMAIN 21336 21377  
FT DOMAIN 21378 21419  
FT DOMAIN 21420 21461  
FT DOMAIN 21462 21503  
FT DOMAIN 21504 21545  
FT DOMAIN 21546 21587  
FT DOMAIN 21588 21629  
FT DOMAIN 21630 21671  
FT DOMAIN 21672 21713  
FT DOMAIN 21714 21755  
FT DOMAIN 21756 21797  
FT DOMAIN 21798 21839  
FT DOMAIN 21840 21881  
FT DOMAIN 21882 21923  
FT DOMAIN 21924 21965  
FT DOMAIN 21966 22007  
FT DOMAIN 22008 22049  
FT DOMAIN 22050 22091  
FT DOMAIN 22092 22133  
FT DOMAIN 22134 22175  
FT DOMAIN 22176 22217  
FT DOMAIN 22218 22259  
FT DOMAIN 22260 22301  
FT DOMAIN 22302 22343  
FT DOMAIN 22344 22385  
FT DOMAIN 22386 22427  
FT DOMAIN 22428 22469  
FT DOMAIN 22470 22511  
FT DOMAIN 22512 22553  
FT DOMAIN 22554 22595  
FT DOMAIN 22596 22637  
FT DOMAIN 22638 22679  
FT DOMAIN 22680 22721  
FT DOMAIN 22722 22763  
FT DOMAIN 22764 22805  
FT DOMAIN 22806 22847  
FT DOMAIN 22848 22889  
FT DOMAIN 22890 22931  
FT DOMAIN 22932 22973  
FT DOMAIN 22974 23015  
FT DOMAIN 23016 23057  
FT DOMAIN 23058 23100  
FT DOMAIN 23101 23142  
FT DOMAIN 23143 23184  
FT DOMAIN 23185 23226  
FT DOMAIN 23227 23268  
FT DOMAIN 23269 23310  
FT DOMAIN 23311 23352  
FT DOMAIN 23353 23394  
FT DOMAIN 23395 23437  
FT DOMAIN 23438 23479  
FT DOMAIN 23480 23521  
FT DOMAIN 23522 23563  
FT DOMAIN 23564 23605  
FT DOMAIN 23606 23647  
FT DOMAIN 23648 23689  
FT DOMAIN 23690 23731  
FT DOMAIN 23732 23773  
FT DOMAIN 23774 23815  
FT DOMAIN 23816 23857  
FT DOMAIN 23858 23899  
FT DOMAIN 23900 23941  
FT DOMAIN 23942 23983  
FT DOMAIN 23984 24025  
FT DOMAIN 24026 24067  
FT DOMAIN 24068 24109  
FT DOMAIN 24110 24151  
FT DOMAIN 24152 24193  
FT DOMAIN 24194 24235  
FT DOMAIN 24236 24277  
FT DOMAIN 24278 24319  
FT DOMAIN 24320 24361  
FT DOMAIN 24362 24403  
FT DOMAIN 24404 24445  
FT DOMAIN 24446 24487  
FT DOMAIN 24488 24529  
FT DOMAIN 24530 24571  
FT DOMAIN 24572 24613  
FT DOMAIN 24614 24655  
FT DOMAIN 24656 24697  
FT DOMAIN 24698 24739  
FT DOMAIN 24740 24781  
FT DOMAIN 24782 24823  
FT DOMAIN 24824 24865  
FT DOMAIN 24866 24907  
FT DOMAIN 24908 24949  
FT DOMAIN 24950 24991  
FT DOMAIN 24992 25033  
FT DOMAIN 25034 25075  
FT DOMAIN 25076 25117  
FT DOMAIN 25118 25159  
FT DOMAIN 25160 25201  
FT DOMAIN 25202 25243  
FT DOMAIN 25244 25285  
FT DOMAIN 25286 25327  
FT DOMAIN 25328 25369  
FT DOMAIN 25370 25411  
FT DOMAIN 25412 25453  
FT DOMAIN 25454 25495  
FT DOMAIN 25496 25537  
FT DOMAIN 25538 25579  
FT DOMAIN 25580 25621  
FT DOMAIN 25622 25663  
FT DOMAIN 25664 25705  
FT DOMAIN 25706 25747  
FT DOMAIN 25748 25789  
FT DOMAIN 25790 25831  
FT DOMAIN 25832 25873  
FT DOMAIN 25874 25915  
FT DOMAIN 25916 25957  
FT DOMAIN 25958 26000  
FT DOMAIN 26001 26042  
FT DOMAIN 26043 26084  
FT DOMAIN 26085 26126  
FT DOMAIN 26127 26168  
FT DOMAIN 26169 26210  
FT DOMAIN 26211 26252  
FT DOMAIN 26253 26294  
FT DOMAIN 26295 26336  
FT DOMAIN 26337 26378  
FT DOMAIN 26379 26420  
FT DOMAIN 26421 26462  
FT DOMAIN 26463 26504  
FT DOMAIN 26505 26546  
FT DOMAIN 26547 26588  
FT DOMAIN 26589 26630  
FT DOMAIN 26631 26672  
FT DOMAIN 26673 26714  
FT DOMAIN 26715 26756  
FT DOMAIN 26757 26798  
FT DOMAIN 26799 26840  
FT DOMAIN 26841 26882  
FT DOMAIN 26883 26923  
FT DOMAIN 26924 26965  
FT DOMAIN 26966 27007  
FT DOMAIN 27008 27049  
FT DOMAIN 27050 27091  
FT DOMAIN 27092 27133  
FT DOMAIN 27134 27175  
FT DOMAIN 27176 27217  
FT DOMAIN 27218 27259  
FT DOMAIN 27260 27301  
FT DOMAIN 27302 27343  
FT DOMAIN 27344 27385  
FT DOMAIN 27386 27427  
FT DOMAIN 27428 27469  
FT DOMAIN 27470 27511  
FT DOMAIN 27512 27553  
FT DOMAIN 27554 27595  
FT DOMAIN 27596 27637  
FT DOMAIN 27638 27679  
FT DOMAIN 27680 27721  
FT DOMAIN 27722 27763  
FT DOMAIN 27764 27805  
FT DOMAIN 27806 27847  
FT DOMAIN 27848 27889  
FT DOMAIN 27890 27931  
FT DOMAIN 27932 27973  
FT DOMAIN 27974 28015  
FT DOMAIN 28016 28057  
FT DOMAIN 28058 28099  
FT DOMAIN 28100 28141  
FT DOMAIN 28142 28183  
FT DOMAIN 28184 28225  
FT DOMAIN 28226 28267  
FT DOMAIN 28268 28309  
FT DOMAIN 28310 28351  
FT DOMAIN 28352 28393  
FT DOMAIN 28394 28435  
FT DOMAIN 28436 28477  
FT DOMAIN 28478 28519  
FT DOMAIN 28520 28561  
FT DOMAIN 28562 28603  
FT DOMAIN 28604 28645  
FT DOMAIN 28646 28687  
FT DOMAIN 28688 28729  
FT DOMAIN 28730 28771  
FT DOMAIN 28772 28813  
FT DOMAIN 28814 28855  
FT DOMAIN 28856 28897  
FT DOMAIN 28898 28939  
FT DOMAIN 28940 28981  
FT DOMAIN 28982 29023  
FT DOMAIN 29024 29065  
FT DOMAIN 29066 29107





```
FT DISULFID 182 BY SIMILARITY.
FT DISULFID 189 BY SIMILARITY.
FT DISULFID 202 BY SIMILARITY.
FT DISULFID 223 BY SIMILARITY.
FT DISULFID 230 BY SIMILARITY.
FT DISULFID 240 BY SIMILARITY.
FT DISULFID 242 BY SIMILARITY.
FT DISULFID 262 BY SIMILARITY.
FT DISULFID 269 BY SIMILARITY.
FT DISULFID 282 BY SIMILARITY.
FT DISULFID 302 BY SIMILARITY.
FT DISULFID 307 BY SIMILARITY.
FT DISULFID 322 BY SIMILARITY.
FT DISULFID 342 BY SIMILARITY.
FT DISULFID 348 BY SIMILARITY.
FT DISULFID 363 BY SIMILARITY.
FT DISULFID 673 BY SIMILARITY.
FT DISULFID 678 BY SIMILARITY.
FT DISULFID 699 BY SIMILARITY.
FT DISULFID 1001 BY SIMILARITY.
FT DISULFID 1006 BY SIMILARITY.
FT DISULFID 1028 BY SIMILARITY.
FT DISULFID 1054 BY SIMILARITY.
FT DISULFID 1063 BY SIMILARITY.
FT DISULFID 1075 BY SIMILARITY.
FT DISULFID 1101 BY SIMILARITY.
FT DISULFID 1108 BY SIMILARITY.
FT DISULFID 1121 BY SIMILARITY.
FT DISULFID 1146 BY SIMILARITY.
FT DISULFID 1153 BY SIMILARITY.
FT DISULFID 1165 BY SIMILARITY.
FT DISULFID 1187 BY SIMILARITY.
FT DISULFID 1194 BY SIMILARITY.
FT DISULFID 1206 BY SIMILARITY.
FT DISULFID 1228 BY SIMILARITY.
FT DISULFID 1235 BY SIMILARITY.
FT DISULFID 1248 BY SIMILARITY.
FT DISULFID 1270 BY SIMILARITY.
FT DISULFID 1277 BY SIMILARITY.
FT DISULFID 1290 BY SIMILARITY.
FT DISULFID 1313 BY SIMILARITY.
FT DISULFID 1320 BY SIMILARITY.
FT DISULFID 1332 BY SIMILARITY.
FT DISULFID 1359 BY SIMILARITY.
FT DISULFID 1366 BY SIMILARITY.
FT DISULFID 1380 BY SIMILARITY.
FT DISULFID 1401 BY SIMILARITY.
FT DISULFID 1408 BY SIMILARITY.
FT DISULFID 1423 BY SIMILARITY.
FT DISULFID 1441 BY SIMILARITY.
FT DISULFID 1447 BY SIMILARITY.
FT DISULFID 1462 BY SIMILARITY.
FT DISULFID 1751 BY SIMILARITY.
FT DISULFID 1756 BY SIMILARITY.
FT DISULFID 1772 BY SIMILARITY.
FT DISULFID 2084 BY SIMILARITY.
FT DISULFID 2091 BY SIMILARITY.
FT DISULFID 2107 BY SIMILARITY.
FT DISULFID 2400 BY SIMILARITY.
FT DISULFID 2411 BY SIMILARITY.
FT DISULFID 2428 BY SIMILARITY.
FT DISULFID 2732 BY SIMILARITY.
FT DISULFID 2739 BY SIMILARITY.
FT DISULFID 2761 BY SIMILARITY.
FT DISULFID 2792 BY SIMILARITY.
FT DISULFID 2800 BY SIMILARITY.
FT DISULFID 2812 BY SIMILARITY.
FT DISULFID 2834 BY SIMILARITY.
FT DISULFID 2841 BY SIMILARITY.

195 BY SIMILARITY.
208 BY SIMILARITY.
238 BY SIMILARITY.
235 BY SIMILARITY.
248 BY SIMILARITY.
257 BY SIMILARITY.
275 BY SIMILARITY.
288 BY SIMILARITY.
297 BY SIMILARITY.
311 BY SIMILARITY.
320 BY SIMILARITY.
336 BY SIMILARITY.
352 BY SIMILARITY.
361 BY SIMILARITY.
367 BY SIMILARITY.
682 BY SIMILARITY.
697 BY SIMILARITY.
711 BY SIMILARITY.
1010 BY SIMILARITY.
1036 BY SIMILARITY.
1042 BY SIMILARITY.
1068 BY SIMILARITY.
1081 BY SIMILARITY.
1095 BY SIMILARITY.
1114 BY SIMILARITY.
1127 BY SIMILARITY.
1138 BY SIMILARITY.
1158 BY SIMILARITY.
1171 BY SIMILARITY.
1182 BY SIMILARITY.
1199 BY SIMILARITY.
1212 BY SIMILARITY.
1223 BY SIMILARITY.
1241 BY SIMILARITY.
1254 BY SIMILARITY.
1263 BY SIMILARITY.
1283 BY SIMILARITY.
1296 BY SIMILARITY.
1307 BY SIMILARITY.
1325 BY SIMILARITY.
1338 BY SIMILARITY.
1350 BY SIMILARITY.
1373 BY SIMILARITY.
1386 BY SIMILARITY.
1396 BY SIMILARITY.
1412 BY SIMILARITY.
1421 BY SIMILARITY.
1435 BY SIMILARITY.
1451 BY SIMILARITY.
1460 BY SIMILARITY.
1475 BY SIMILARITY.
1760 BY SIMILARITY.
1770 BY SIMILARITY.
1785 BY SIMILARITY.
2095 BY SIMILARITY.
2105 BY SIMILARITY.
2119 BY SIMILARITY.
2415 BY SIMILARITY.
2426 BY SIMILARITY.
2438 BY SIMILARITY.
2743 BY SIMILARITY.
2759 BY SIMILARITY.
2779 BY SIMILARITY.
2805 BY SIMILARITY.
2818 BY SIMILARITY.
2829 BY SIMILARITY.
2846 BY SIMILARITY.
2859 BY SIMILARITY.

Query Match 19.0%; Score 92.5; DB 1; Length 4753;
Best Local Similarity 25.0%; Pred. No. 0.69;
Matches 28; Conservative 11; Mismatches 32; Indels 41; Gaps 5;

OY 7 GENEKDYSCGSKC-----DKCKYDGVVEEDDEPNVPC----- 41
DB 1174 GSDKLEKCGNATCAANQFSCANGRCIPYWLCDGNDG- YDGTDEKRCPPVQCSALQ 1232
OY 42 -----LVRVCHQDCVCBEGFYRNKDDKC-VSAEDCELDNMDFIYPG 81
DB 1233 FRCANGRCQVPLRNHCDGQSDCEDG---SDEDSCAVTAESCTPDQFKCVSSG 1281

RESULT 10
VWF_CANFA STANDARD; PRT; 2813 AA.
AC Q28295; Q28311; Q9T5I4;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Von Willebrand factor precursor (VWF).
GN VWF OR F8VWF.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Stoy S.J., Shibuya H., Nonneman D.J., Holzhauer J., Mohammed I.H.,
RA Johnson G.S.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA Montgomery R.R., Fahs S., Montgomery M.W.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA Venta P.J., Li J., Yuzbasiyan-Gurkan V., Brewer G.J., Schall W.D.;
RT "Complete sequence of the structural gene for canine von Willebrand
RT factor and identification of a mutation causing Scottish terrier von
RT Willebrand's disease.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE OF 1234-1669 FROM N.A.
RC TISSUE=Blood;
RA Mancuso D.J., Christopherson P.A., Kroner P.A., Montgomery R.R.;
RT "The canine von Willebrand factor gene: sequence and expression of
RT a region encoding the glycoprotein Ib/IX binding domain.";
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Important in the maintenance of homeostasis, it
CC participates in platelet-vessel wall interactions by forming a
CC noncovalent complex with coagulation factor VIII at the site of
CC vascular injury (By similarity).
CC -!- SUBUNIT: Multimeric (By similarity).
CC -!- TISSUE SPECIFICITY: Blood.
CC -!- PTM: ALL CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN OR
CC INTERCHAIN DISULFIDE BONDS (BY SIMILARITY).
CC -!- SIMILARITY: Contains 3 VWF domains.
CC -!- SIMILARITY: Contains 3 VWF domains.
CC -!- SIMILARITY: Contains 4 VWF domains.
CC -!- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC -!- SIMILARITY: SOME, TO SILKWORM HEMOCYTIN.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; L76227; AAB05549.1; --
DR EMBL; L16903; AAA30903.1; --
DR EMBL; AF099154; AAD04919.1; --
DR EMBL; U68246; AAB93766.1; --
DR HSSP; P04275; LAUQ.
DR InterPro; IPR006208; Cys_knot.
```

DR InterPro; IPR006207; Cys knot C.  
 DR InterPro; IPR008139; SaposinB\_  
 DR InterPro; IPR002919; TIL Cysrich.  
 DR InterPro; IPR002035; VWF A.  
 DR InterPro; IPR001007; VWF C.  
 DR InterPro; IPR001846; VWF D.  
 DR Pfam; PF00007; Cys knot; 1.  
 DR Pfam; PF01826; TIL; 4.  
 DR Pfam; PF00092; vws; 3.  
 DR Pfam; PF00093; vwc; 3.  
 DR Pfam; PF00094; vwd; 4.  
 DR PRINTS; PR00453; VWFADOMAIN.  
 DR SMART; SM00041; CT; 1.  
 DR SMART; SM00327; VWA; 3.  
 DR SMART; SM00214; VWC; 5.  
 DR SMART; SM00216; VWD; 4.  
 DR PROSITE; PS01185; CTCK 1; 1.  
 DR PROSITE; PS01225; CTCK 2; 1.  
 DR PROSITE; PS0234; VWFA; 3.  
 DR PROSITE; PS01208; VWFC 1; 3.  
 DR PROSITE; PS01184; VWFC 2; 3.  
 KW Blood coagulation; Platelet; Glycoprotein; Extracellular matrix;  
 KW Plasma; Hemostasis; Repeat; Cell adhesion; Signal.  
 FT SIGNAL 1 22  
 FT PROPEP 23 763  
 FT CHAIN 764 2813  
 FT DOMAIN 35 179  
 FT DOMAIN 388 541  
 FT DOMAIN 764 787  
 FT DOMAIN 788 833  
 FT DOMAIN 826 853  
 FT DOMAIN 867 1013  
 FT DOMAIN 1277 1453  
 FT DOMAIN 1498 1665  
 FT DOMAIN 1691 1871  
 FT DOMAIN 1950 2102  
 FT DOMAIN 2216 2261  
 FT DOMAIN 2255 2328  
 FT DOMAIN 2429 2495  
 FT DOMAIN 2580 2645  
 FT DOMAIN 2724 2812  
 FT SITE 531 533  
 FT SITE 698 700  
 FT SITE 2507 2509  
 FT SITE 767 808  
 FT DISULFID 776 804  
 FT DISULFID 810 821  
 FT DISULFID 867 996  
 FT DISULFID 889 1031  
 FT DISULFID 898 993  
 FT DISULFID 914 921  
 FT DISULFID 1060 1084  
 FT DISULFID 1071 1111  
 FT DISULFID 1089 1091  
 FT DISULFID 1153 1165  
 FT DISULFID 1149 1169  
 FT DISULFID 1126 1130  
 FT DISULFID 1196 1199  
 FT DISULFID 1234 1237  
 FT DISULFID 1272 1458  
 FT DISULFID 1669 1670  
 FT DISULFID 1686 1872  
 FT DISULFID 1879 1904  
 FT DISULFID 1899 1940  
 FT DISULFID 1972 2123  
 FT DISULFID 1950 2085  
 FT DISULFID 1927 2088  
 FT DISULFID 1993 2001  
 FT DISULFID 2724 2774  
 FT DISULFID 2739 2788  
 FT DISULFID 2750 2804  
 FT DISULFID 2754 2806  
 FT DISULFID 2754 2811

FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 211 211 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 666 666 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 857 857 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1231 1231 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1515 1515 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1574 1574 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 2223 2223 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 2290 2290 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 2357 2357 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 2400 2400 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 2546 2546 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 2585 2585 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 2790 2790 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 55 55 G -> E (IN REF. 2).  
 FT CONFLICT 70 70 V -> I (IN REF. 3).  
 FT CONFLICT 266 266 A -> G (IN REF. 2).  
 FT CONFLICT 280 280 I -> V (IN REF. 2).  
 FT CONFLICT 409 411 VCH -> ICQ (IN REF. 2).  
 FT CONFLICT 994 994 G -> A (IN REF. 1).  
 FT CONFLICT 1021 1021 F -> L (IN REF. 2).  
 FT CONFLICT 2381 2381 L -> P (IN REF. 2).  
 FT CONFLICT 2406 2406 P -> L (IN REF. 2).  
 SQ SEQUENCE 2813 AA, 309716 MW, 5DF93E1E5E72F80C CRC64;  
 Query Match 18.0%; Score 87.5; DB 1; Length 2813;  
 Best Local Similarity 28.4%; Pred. No. 1.2;  
 Matches 19; Conservative 9; Mismatches 30; Indels 9; Gaps 2;  
 QY 4 MQGENEKYDSCGKCDKKYDGVVEEDDEPNVPCLVRVCHQDCVCEGPFYRKDDK 63  
 Db 550 LSPQGGVYLQGT-PCNMTCSRSLSPEDCNE-----VLEGCFCPPGLDREGD 700  
 QY 64 CVSAEDC 70  
 Db 701 CVPKACQ 707  
 RESULT 11  
 ID A62F DROME STANDARD; PRT; 115 AA.  
 AC O46202;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Accessory gland protein Acp62F precursor.  
 GN ACP62F OR CGI262.  
 OS Drosophila melanogaster (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OC NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.  
 RC STRAIN=Carton-S.A.; TISSUE=Male accessory gland;  
 RX MEDLINE=98135120; PubMed=9474779;  
 RA Wolfner M.F., Harada H.A., Bertram M.J., Stelick T.J., Kraus K.W.,  
 RA Kalb J.M., Lung Y.O., Neubaum D.M., Park M., Tram U.K.;  
 RA "New genes for male accessory gland proteins in Drosophila  
 RT melanogaster.";  
 RL Insect Biochem. Mol. Biol. 27:825-834 (1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlike C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Garg N.S., Galbart W.M., Glasner K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harlis N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.R., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
RA "The genome sequence of *Drosophila melanogaster*,"  
RL Science 287:2185-2195(2000).  
[3]  
RP SEQUENCE OF 7-111 FROM N.A.  
RC STRAIN=ZIM62H-12C, ZIM62H-16C, ZIM62H-28C, ZIM62H-30C, ZIM62H-34C,  
RC ZIM62H-5C, ZIM62H-10C, ZIM62H-17C, ZIM62H-18C, and ZIM62H-53C;  
RX MEDLINE=20556153; PubMed=11102381;  
RA Begun D.J., Whitley P., Todd B.L., Waldrip-Dail H.M., Clark A.G.,  
RT "Molecular population genetics of male accessory gland proteins in  
RT *Drosophila*,"  
RL Genetics 156:1879-1898(2000).  
CC -!- FUNCTION: RESPONSIBLE FOR PHYSIOLOGICAL AND BEHAVIORAL CHANGES IN  
CC MATED FEMALE FLIES. MAY CONTRIBUTE TO THE TOXICITY OF SEMINAL  
CC FLUID AND THE DECREASED LIFE-SPAN OF MATED FEMALES. MAY ALSO  
CC AFFECT NEUROMUSCULAR EVENTS AFTER MATING CONCERNING SPERM STORAGE  
CC AND EGG RELEASE.  
CC -!- SUBCELLULAR LOCATION: Secreted (Probable).  
CC -!- TISSUE SPECIFICITY: Seminal fluid.  
CC -!- SIMILARITY: SOME, TO P.NIGRIVENTER TX2-6.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC -----  
DR EMBL; U85763; AAB96387.1; -  
DR EMBL; AE003475; AAF47683.1; -  
DR EMBL; AY010608; AAG35367.1; -  
DR EMBL; AY010609; AAG35368.1; -  
DR EMBL; AY010610; AAG35369.1; -  
DR EMBL; AY010611; AAG35370.1; -  
DR EMBL; AY010612; AAG35371.1; -  
DR EMBL; AY010613; AAG35372.1; -  
DR EMBL; AY010614; AAG35373.1; -  
DR EMBL; AY010615; AAG35374.1; -  
DR EMBL; AY010616; AAG35375.1; -  
DR EMBL; AY010617; AAG35376.1; -  
DR FlyBase; FBgn020509; ACP62F.  
DR GO; GO:0004867; F:serine protease inhibitor activity; IDA.  
DR GO; GO:0008340; P:serine proteolysis of adult life span; NAS.  
DR InterPro; IPR002919; TIL\_Cysrich.

DR Pfam; PF01826; TIL; 1.  
KW Behavior; Signal.  
FT SIGNAL 1 24 POTENTIAL.  
FT CHAIN 25 115 ACCESSORY GLAND PROTEIN ACP62F.  
FT DOMAIN 34 88 TIL.  
SQ SEQUENCE 115 AA; 12570 MW; 4326AA6F6C32291D CRC64;  
Query Match 17.9%; Score 87; DB 1; Length 115;  
Best Local Similarity 31.0%; Pred. No. 0.054;  
Matches 22; Conservative 8; Mismatches 29; Indels 12; Gaps 4;  
QY 4 MQCCENEKYSCGSKGCKKCKGVDEEDDEENVPCLVRVCHQDCVCEGFYRN-KDD 62  
Db 32 VDCVTANGTQTEC-PVACPFEICVSG-----NGFC-VKMGAPCVCKPGYVNERIP 80  
QY 63 KCVSAEDCELD 73  
Db 81 ACVLRSDCPKD 91  
RESULT 12  
LRP2\_HUMAN STANDARD; PRT; 4655 AA.  
ID LRP2\_HUMAN  
AC P98164; O00711; Q16215;  
DT 01-OCT-1996 (Rel. 34, Created)  
DI 15-DEC-1998 (Rel. 37, Last sequence update)  
DI 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Low-density lipoprotein receptor-related protein 2 precursor (Megalin)  
DE (Glycoprotein 330) (gp330).  
GN LRP2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RX MEDLINE=96305376; PubMed=8706697;  
RA Hjaelm G., Murray E., Crumley G., Harazin W., Lundgren S., Onyango I.,  
RA Ek B., Larsson M., Juhlin C., Hellman P., Davis H., Akerstrom G.,  
RA Rask L., Morse B.,  
RA "Cloning and sequencing of human gp330, a Ca(2+)-binding receptor  
RT with potential intracellular signaling properties,"  
RL Eur. J. Biochem. 239:132-137(1996).  
[2]  
RP SEQUENCE OF 2705-4453 FROM N.A.  
RC TISSUE=Kidney;  
RA Knaak C., Argraves W.S.;  
RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.  
[3]  
RP SEQUENCE OF 3833-4453 FROM N.A.  
RC TISSUE=Kidney;  
RX MEDLINE=95048397; PubMed=7959795;  
RA Korenberg J.R., Argraves K.M., Chen X.N., Tran H.,  
RA Strickland D.K., Akerstrom G., Rask L.;  
RT "Chromosomal localization of human genes for the LDL receptor family  
RT member glycoprotein 330 (LRP2) and its associated protein RAP  
RT (LRPAP1).";  
RL Genomics 22:88-93(1994).  
[4]  
RP SEQUENCE OF 4139-4406 FROM N.A.  
RX MEDLINE=94244704; PubMed=8187828;  
RA Lundgren S., Hjaelm G., Hellman P., Ek B., Juhlin C., Rastad J.,  
RA Klarskog L., Akerstrom G., Rask L.;  
RT "A protein involved in calcium sensing of the human parathyroid and  
RT placental cytotrophoblast cells belongs to the LDL-receptor protein  
RT superfamily,"  
RL Exp. Cell Res. 212:344-350(1994).  
[5]  
RP FUNCTION.  
RX MEDLINE=95286588; PubMed=7768901;  
RA Kounnas M.Z., Loukinova E.B., Stefansson S., Harmony J.A.K.,  
RA Brewer B.H., Strickland D.K., Argraves W.S.;

RT "Identification of glycoprotein 330 as an endocytic receptor for  
 RL apolipoprotein J/clusterin.";  
 J. Biol. Chem. 270:13070-13075(1995).  
 CC -I- FUNCTION: Binds specifically clusterin with high affinity. But  
 also ligands in common with other family members: plasminogen,  
 extracellular matrix components, plasminogen activator-plasminogen  
 activator inhibitor type I complex, apolipoprotein E-enriched  
 beta-VLDL, lipoprotein lipase, lactoferrin and calcium.  
 CC -I- FUNCTION: Receptor-mediated uptake of polybasic drugs such as  
 apoferritin, aminoglycosides and polymyxin B (By similarity).  
 CC -I- FUNCTION: May participate in regulation of parathyroid-hormone and  
 para-thyroid-hormone-related protein release.  
 CC -I- SUBUNIT: Forms a multimeric complex together with a receptor-  
 associated protein (RAP). Binds to ankyrin-repeat family A protein  
 2 (ANKRA2) (By similarity).  
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -I- TISSUE SPECIFICITY: Absorptive epithelia, including renal  
 proximal tubules.  
 CC -I- SIMILARITY: Contains 36 LDL-receptor class A domains.  
 CC -I- SIMILARITY: Contains 37 LDL-receptor class B domains.  
 CC -I- SIMILARITY: Contains 17 EGF-like domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U33837; AAB41649.1; -;  
 DR EMBL: U04441; AAB02882.1; -;  
 DR EMBL: S73145; AAB30825.1; -;  
 DR PIR: I53413; I53413.  
 DR HSP: Q07954; ICR8.  
 DR Genew: HGNC:6694; LRP2.  
 DR MIM: 600073; -;  
 DR GO: GO:0005764; C:cytosome; TAS.  
 DR GO: GO:0006629; P:lipid metabolism; TAS.  
 DR GO: GO:0006486; P:protein amino acid glycosylation; TAS.  
 DR GO: GO:0006898; P:receptor mediated endocytosis; TAS.  
 DR InterPro: IPR001012; Asx\_hydroxyl\_5.  
 DR InterPro: IPR001881; EGF\_Ca.  
 DR InterPro: IPR006209; EGF\_Like.  
 DR InterPro: IPR002172; LDL\_receptor\_A.  
 DR InterPro: IPR000033; LDL\_receptor\_rep.  
 DR Pfam: PF00008; EGF\_10.  
 DR Pfam: PF00057; ldl\_recept\_a; 36.  
 DR Pfam: PF00058; ldl\_recept\_b; 37.  
 DR SMART: SM00179; EGF\_CA; 3.  
 DR SMART: SM00192; LDLa; 36.  
 DR SMART: SM00135; LV; 33.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; 4.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 9.  
 DR PROSITE: PS00026; EGF\_3; 6.  
 DR PROSITE: PS01187; EGF\_CA; 3.  
 DR PROSITE: PS01209; LDLRA\_1; 31.  
 DR PROSITE: PS00068; LDLRA\_2; 36.  
 KW Glycoprotein; Repeat; Endocytosis; Coated pits; Transmembrane;  
 KW Receptor; EGF-like domain; Signal; SH3-binding; Polymorphism.  
 FT SIGNAL 1 25  
 FT CHAIN 26 4655 LOW-DENSITY LIPOPROTEIN RECEPTOR-RELATED  
 FT PROTEIN 2.  
 FT DOMAIN 26 4423 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 4424 4446 POTENTIAL.  
 FT DOMAIN 4447 4655 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 26 64 LDL-RECEPTOR CLASS A 1.  
 FT DOMAIN 65 105 LDL-RECEPTOR CLASS A 2.  
 FT DOMAIN 106 144 LDL-RECEPTOR CLASS A 3.  
 FT DOMAIN 145 181 LDL-RECEPTOR CLASS A 4.  
 FT DOMAIN 182 219 LDL-RECEPTOR CLASS A 5.  
 FT  
 220 DOMAIN  
 264 DOMAIN  
 309 DOMAIN  
 348 DOMAIN  
 436 DOMAIN  
 479 DOMAIN  
 522 DOMAIN  
 569 DOMAIN  
 613 DOMAIN  
 659 DOMAIN  
 753 DOMAIN  
 794 DOMAIN  
 836 DOMAIN  
 880 DOMAIN  
 924 DOMAIN  
 970 DOMAIN  
 1024 DOMAIN  
 1065 DOMAIN  
 1107 DOMAIN  
 1145 DOMAIN  
 1185 DOMAIN  
 1224 DOMAIN  
 1268 DOMAIN  
 1307 DOMAIN  
 1350 DOMAIN  
 1389 DOMAIN  
 1429 DOMAIN  
 1478 DOMAIN  
 1521 DOMAIN  
 1566 DOMAIN  
 1610 DOMAIN  
 1655 DOMAIN  
 1700 DOMAIN  
 1790 DOMAIN  
 1831 DOMAIN  
 1881 DOMAIN  
 1929 DOMAIN  
 1983 DOMAIN  
 2018 DOMAIN  
 2059 DOMAIN  
 2107 DOMAIN  
 2155 DOMAIN  
 2200 DOMAIN  
 2244 DOMAIN  
 2288 DOMAIN  
 2331 DOMAIN  
 2342 DOMAIN  
 2431 DOMAIN  
 2476 DOMAIN  
 2517 DOMAIN  
 2551 DOMAIN  
 2593 DOMAIN  
 2603 DOMAIN  
 2646 DOMAIN  
 2693 DOMAIN  
 2698 DOMAIN  
 2739 DOMAIN  
 2778 DOMAIN  
 2819 DOMAIN  
 2861 DOMAIN  
 2901 DOMAIN  
 2945 DOMAIN  
 2991 DOMAIN  
 3029 DOMAIN  
 3070 DOMAIN  
 3073 DOMAIN  
 3111 DOMAIN  
 3151 DOMAIN  
 3192 DOMAIN  
 3239 DOMAIN  
 3282 DOMAIN  
 3333 DOMAIN  
 3377 DOMAIN  
 3418 DOMAIN  
 3459 DOMAIN  
 3465 DOMAIN  
 3510 DOMAIN  
 3551 DOMAIN  
 3592 DOMAIN  
 3633 DOMAIN  
 3676 DOMAIN  
 3717 DOMAIN  
 3756 DOMAIN  
 258 LDL-RECEPTOR CLASS A 6.  
 308 LDL-RECEPTOR CLASS A 7.  
 347 EGF-LIKE 1.  
 386 EGF-LIKE 2.  
 477 LDL-RECEPTOR CLASS B 1.  
 520 LDL-RECEPTOR CLASS B 2.  
 567 LDL-RECEPTOR CLASS B 3.  
 612 LDL-RECEPTOR CLASS B 4.  
 653 LDL-RECEPTOR CLASS B 5.  
 694 EGF-LIKE 3.  
 734 LDL-RECEPTOR CLASS B 6.  
 794 LDL-RECEPTOR CLASS B 7.  
 836 LDL-RECEPTOR CLASS B 8.  
 880 LDL-RECEPTOR CLASS B 9.  
 924 EGF-LIKE 4.  
 970 LDL-RECEPTOR CLASS A 8.  
 1024 LDL-RECEPTOR CLASS A 9.  
 1065 LDL-RECEPTOR CLASS A 10.  
 1107 LDL-RECEPTOR CLASS A 11.  
 1145 LDL-RECEPTOR CLASS A 12.  
 1185 LDL-RECEPTOR CLASS A 13.  
 1224 LDL-RECEPTOR CLASS A 14.  
 1268 LDL-RECEPTOR CLASS A 15.  
 1307 EGF-LIKE 5.  
 1350 LDL-RECEPTOR CLASS B 10.  
 1389 LDL-RECEPTOR CLASS B 11.  
 1429 LDL-RECEPTOR CLASS B 12.  
 1478 LDL-RECEPTOR CLASS B 13.  
 1521 LDL-RECEPTOR CLASS B 14.  
 1566 EGF-LIKE 6.  
 1610 CALCIUM-BINDING (POTENTIAL).  
 1655 LDL-RECEPTOR CLASS B 15.  
 1700 LDL-RECEPTOR CLASS B 16.  
 1790 LDL-RECEPTOR CLASS B 17.  
 1831 LDL-RECEPTOR CLASS B 18.  
 1881 LDL-RECEPTOR CLASS B 19.  
 1929 EGF-LIKE 8.  
 1983 LDL-RECEPTOR CLASS B 20.  
 2018 LDL-RECEPTOR CLASS B 21.  
 2059 LDL-RECEPTOR CLASS B 22.  
 2107 LDL-RECEPTOR CLASS B 23.  
 2155 LDL-RECEPTOR CLASS B 24.  
 2200 EGF-LIKE 9.  
 2244 LDL-RECEPTOR CLASS B 25.  
 2288 LDL-RECEPTOR CLASS B 26.  
 2331 LDL-RECEPTOR CLASS B 27.  
 2342 LDL-RECEPTOR CLASS B 28.  
 2431 LDL-RECEPTOR CLASS B 29.  
 2476 EGF-LIKE 10.  
 2517 LDL-RECEPTOR CLASS A 16.  
 2551 LDL-RECEPTOR CLASS A 17.  
 2593 LDL-RECEPTOR CLASS A 18.  
 2603 LDL-RECEPTOR CLASS A 19.  
 2646 LDL-RECEPTOR CLASS A 20.  
 2693 LDL-RECEPTOR CLASS A 21.  
 2698 LDL-RECEPTOR CLASS A 22.  
 2739 LDL-RECEPTOR CLASS A 23.  
 2778 LDL-RECEPTOR CLASS A 24.  
 2819 LDL-RECEPTOR CLASS A 25.  
 2861 EGF-LIKE 11.  
 2901 CALCIUM-BINDING (POTENTIAL).  
 2945 EGF-LIKE 12.  
 2991 LDL-RECEPTOR CLASS B 30.  
 3029 LDL-RECEPTOR CLASS B 31.  
 3070 LDL-RECEPTOR CLASS B 32.  
 3073 LDL-RECEPTOR CLASS B 33.  
 3111 LDL-RECEPTOR CLASS B 34.  
 3151 EGF-LIKE 13.  
 3192 LDL-RECEPTOR CLASS A 26.  
 3239 LDL-RECEPTOR CLASS A 27.  
 3282 LDL-RECEPTOR CLASS A 28.  
 3333 LDL-RECEPTOR CLASS A 29.  
 3377 LDL-RECEPTOR CLASS A 30.  
 3418 LDL-RECEPTOR CLASS A 31.  
 3459  
 3465  
 3510  
 3551  
 3592  
 3633  
 3676  
 3717  
 3756

```
FT DOMAIN 3757 3795 LDL-RECEPTOR CLASS A 32.
FT DOMAIN 3796 3834 LDL-RECEPTOR CLASS A 33.
FT DOMAIN 3840 3880 LDL-RECEPTOR CLASS A 34.

Query Match 17.8%; Score 86.5; DB 1; Length 4655;
Best Local Similarity 26.1%; Pred. No. 2.6; Indels 21; Gaps 5;
Matches 23; Conservative 14; Mismatches 30; Indels 21; Gaps 5;

QY 6 CGENKDYSCGKCEC-----DKKC-----KYDGEDEDEEPNVPCLVRVCHQD 49
Db 3745 CGDNDSEENCAPRECTESFRVCNQICPSRWICDHNDGNSDSR---DCMETCHPE 3801

QY 50 CV-CGEGFYRNKDDKCVSAEDCELDNMD 76
Db 3802 YFQCTSGCHVHSELKCDGSGADC-LDASD 3828

RESULT 13
ID ICE2_ASCSU STANDARD; PRT; 65 AA.
AC P07852;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Chymotrypsin/elastase isoforms 2 to 5.
OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC Ascarididae; Ascaris.
OX NCBI_TaxID=6253;
RN [1]
RP MEDLINE=84255715; PubMed=6564898;
RA Babin D.R., Pearsall R.J., Coos S.M.;
RT "The inhibitors of chymotrypsin/elastase from Ascaris
RT lumbricoides: the primary structure."
RL Arch. Biochem. Biophys. 232:143-161(1984).
CC 1- FUNCTION: Defend the organism against the host's proteinases.
CC 1- SUBCELLULAR LOCATION: Secreted.
CC 1- SIMILARITY: Contains 1 TIL (Trypsin inhibitory-like) domain.
DR HSSP; P07851; LEAI.
DR InterPro; IPR002919; TIL_Cysrich.
DR Pfam; PF01826; TIL; 1.
KW Serine protease inhibitor.
FT DOMAIN 4 59 TIL.
FT DISULFID 4 37 BY SIMILARITY.
FT DISULFID 13 32 BY SIMILARITY.
FT DISULFID 16 28 BY SIMILARITY.
FT DISULFID 20 59 BY SIMILARITY.
FT DISULFID 39 53 BY SIMILARITY.
FT ACT SITE 30 31 REACTIVE BOND.
FT VARIANT 25 25 K -> N (IN INHIBITOR 2 AND 4).
FT VARIANT 40 40 T -> S (IN INHIBITOR 2 AND 4).
FT VARIANT 64 65 MISSING (IN INHIBITOR 2).
FT VARIANT 65 65 K -> E (IN INHIBITOR 3).
FT VARIANT 65 65 K -> R.
SQ SEQUENCE 65 AA; 7241 MW; B4E51CA166EA4BE3 CRC64;

Query Match 17.6%; Score 85.5; DB 1; Length 65;
Best Local Similarity 28.4%; Pred. No. 0.042;
Matches 19; Conservative 7; Mismatches 28; Indels 13; Gaps 3;

QY 6 CGENKDYSCGKCECCKKCYDGEDEDEEPNVPCLVRVCHQDVCVE--EGFYRNKDDK 63
Db 4 CGKNEVTECTG--CELKC-----GQDKTPCALMCRPPSCBCTPGRMRRTHDGK 52

QY 64 CVSAEDC 70
Db 53 CVPVSEC 59

RESULT 14
LDVR_CHKCK STANDARD; PRT; 863 AA.
ID _LDVR_CHKCK
```

```
AC P98165;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Very low-density lipoprotein receptor precursor (VLDL receptor)
DE (Vitellogenin receptor) (VTG receptor).
GN VLDLR OR VTGR. (Chicken).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White leghorn; TISSUE=Ovary;
RX MEDLINE=95045409; PubMed=7957081;
RA Bujo H., Hermann M., Kaderli M.O., Jacobsen L., Sugawara S.,
RA Nimpf J., Yamamoto T., Schneider W.J.;
RT "Chicken oocyte growth is mediated by an eight ligand binding repeat
RT member of the LDL receptor family."
RL EMBO J. 13:5165-5175(1994).
RN [2]
RP SEQUENCE OF 510-518; 546-554 AND 819-827.
RC STRAIN=White leghorn; TISSUE=Placental membrane;
RX MEDLINE=92011638; PubMed=1655760;
RA Barber D.L., Sanders E.J., Abersold R., Schneider W.J.;
RT "The receptor for yolk lipoprotein deposition in the chicken oocyte."
RL J. Biol. Chem. 266:18761-18770(1991).
CC 1- FUNCTION: Binds VLDL and transports it into cells by endocytosis.
CC In order to be internalized, the receptor-ligand complexes must
CC first cluster into clathrin-coated pits. Binding to Reelin induces
CC tyrosine phosphorylation of Dab1 and modulation of Tau
CC phosphorylation (By similarity).
CC 1- SUBUNIT: Binds to the extracellular matrix protein Reelin (By
CC similarity).
CC 1- SUBCELLULAR LOCATION: Type I membrane protein.
CC 1- TISSUE SPECIFICITY: Abundant in oocytes; much less in heart and
CC skeletal muscle.
CC 1- SIMILARITY: Contains 8 LDL-receptor class A domains.
CC 1- SIMILARITY: Contains 6 LDL-receptor class B domains.
CC 1- SIMILARITY: Contains 3 EGF-like domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X80207; CRA56505.1; -.
DR PIR; S51789; S51789.
DR HSSP; P01130; 1AJJ.
DR InterPro; IPR000152; Asx_hydroxyl_S.
DR InterPro; IPR001881; EGF_Ga.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; Ldl_receptor_rep.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF00057; ldl_recept_a; 8.
DR Pfam; PF00058; ldl_recept_b; 5.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00192; LDLA; 8.
DR SMART; SM00135; LY; 5.
DR PROSITE; PS00010; ASX_HYDROXYL; 2.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS01187; EGF_CA; 2.
DR PROSITE; PS01209; LDLA_1; 8.
DR PROSITE; PS00068; LDLRA_2; 8.
KW Glycoprotein; VLDL; Cholesterol metabolism; Lipid transport;
```

KW Endocytosis; Coated pits; Transmembrane; Receptor; Signal;  
 KW EGF-like domain; Repeat.  
 FT SIGNAL 1 43 POTENTIAL.  
 FT CHAIN 44 863 VERY LOW-DENSITY LIPOPROTEIN RECEPTOR.  
 FT DOMAIN 44 785 EXTRACELLULAR (POTENTIAL).  
 FT TRANSEM 786 809 POTENTIAL.  
 FT DOMAIN 810 863 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 49 87 LDL-RECEPTOR CLASS A 1.  
 FT DOMAIN 88 128 LDL-RECEPTOR CLASS A 2.  
 FT DOMAIN 129 169 LDL-RECEPTOR CLASS A 3.  
 FT DOMAIN 170 208 LDL-RECEPTOR CLASS A 4.  
 FT DOMAIN 209 249 LDL-RECEPTOR CLASS A 5.  
 FT DOMAIN 255 293 LDL-RECEPTOR CLASS A 6.  
 FT DOMAIN 294 332 LDL-RECEPTOR CLASS A 7.  
 FT DOMAIN 334 373 LDL-RECEPTOR CLASS A 8.  
 FT DOMAIN 374 413 EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 414 453 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).  
 FT REPEAT 457 498 LDL-RECEPTOR CLASS B 1.  
 FT REPEAT 499 544 LDL-RECEPTOR CLASS B 2.  
 FT REPEAT 545 587 LDL-RECEPTOR CLASS B 3.  
 FT REPEAT 588 631 LDL-RECEPTOR CLASS B 4.  
 FT REPEAT 632 674 LDL-RECEPTOR CLASS B 5.  
 FT REPEAT 675 716 LDL-RECEPTOR CLASS B 6.  
 FT DOMAIN 722 770 EGF-LIKE 3.  
 FT SITE 822 827 ENDOCYTOSIS SIGNAL (POTENTIAL).  
 FT DISULFID 51 63 BY SIMILARITY.  
 FT DISULFID 58 76 BY SIMILARITY.  
 FT DISULFID 70 85 BY SIMILARITY.  
 FT DISULFID 90 102 BY SIMILARITY.  
 FT DISULFID 97 115 BY SIMILARITY.  
 FT DISULFID 109 126 BY SIMILARITY.  
 FT DISULFID 131 145 BY SIMILARITY.  
 FT DISULFID 138 158 BY SIMILARITY.  
 FT DISULFID 152 167 BY SIMILARITY.  
 FT DISULFID 172 184 BY SIMILARITY.  
 FT DISULFID 179 197 BY SIMILARITY.  
 FT DISULFID 191 206 BY SIMILARITY.  
 FT DISULFID 211 223 BY SIMILARITY.  
 FT DISULFID 218 236 BY SIMILARITY.  
 FT DISULFID 230 247 BY SIMILARITY.  
 FT DISULFID 257 269 BY SIMILARITY.  
 FT DISULFID 264 282 BY SIMILARITY.  
 FT DISULFID 276 291 BY SIMILARITY.  
 FT DISULFID 296 308 BY SIMILARITY.  
 FT DISULFID 303 321 BY SIMILARITY.  
 FT DISULFID 315 330 BY SIMILARITY.  
 FT DISULFID 336 349 BY SIMILARITY.  
 FT DISULFID 344 362 BY SIMILARITY.  
 FT DISULFID 356 373 BY SIMILARITY.  
 FT DISULFID 378 389 BY SIMILARITY.  
 FT DISULFID 385 398 BY SIMILARITY.  
 FT DISULFID 400 412 BY SIMILARITY.  
 FT DISULFID 418 428 BY SIMILARITY.  
 FT DISULFID 424 437 BY SIMILARITY.  
 FT DISULFID 439 452 BY SIMILARITY.  
 FT DISULFID 726 739 BY SIMILARITY.  
 FT DISULFID 735 754 BY SIMILARITY.  
 FT DISULFID 756 769 BY SIMILARITY.  
 FT CARBOHYD 169 169 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CARBOHYD 773 773 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 863 AA; 94904 MW; 0672A8748F9A2245 CRC64; 17.6%; Score 85.5; DB 1; Length 863;  
 Best Local Similarity 33.3%; Pred. No. 0.58;  
 Matches 25; Conservative 9; Mismatches 36; Indels 5; Gaps 4;  
 QY 4 MCGENEKDYSCGKFC-DKKCKYDCVEDEEENPVCLVCHQDCV-CEEGYRNDK 61  
 Db 255 VKSTSEV--QCSGGCIHKWRCDDGDDPKDGSDEINCPSTCPDQPRCEDGNCIRGS 312  
 QY 62 DKCVSREDCELDNMD 76  
 Db 313 RQNGVGRDC-LDGTD 326

RESULT 15  
 LDVR HUMAN  
 ID LDVR HUMAN STANDARD; PRT; 873 AA.  
 AC P98155;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Very low-density lipoprotein receptor precursor (VLDL receptor).  
 GN VLDLR.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skeletal muscle;  
 RX MEDLINE=94174378; PubMed=8128315;  
 RA Garvels M.E., Caird M., Britt D., Jackson C.L., Patterson D.,  
 RA Strauss J.F.;  
 RT "Cloning of a cDNA encoding a putative human very low density  
 RT lipoprotein/apolipoprotein E receptor and assignment of the gene to  
 RT chromosome 9pter-p23.";  
 RL Sorat. Cell Mol. Genet. 19:557-569(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Heart;  
 RX MEDLINE=94348496; PubMed=8069294;  
 RA Webb J.C., Patel D.D., Jones M.D., Knight B.L., Soutar A.K.;  
 RT "Characterization and tissue-specific expression of the human 'very  
 RT low density lipoprotein (VLDL) receptor' mRNA.";  
 RL Hum. Mol. Genet. 3:531-537(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94124575; PubMed=8294473;  
 RA Sakai J., Hoshino A., Takahashi S., Miura Y., Ishii H., Suzuki H.,  
 RA Kawarabayashi Y., Yamamoto T.;  
 RT "Structure, chromosome location, and expression of the human very low  
 RT density lipoprotein receptor gene.";  
 RL J. Biol. Chem. 269:2173-2182(1994).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Heart;  
 RX MEDLINE=94292216; PubMed=8020981;  
 RA Oka K., Tsung K.W., Sullivan M., Lindsay E., Baldini A., Chan L.;  
 RT "Human very-low-density lipoprotein receptor complementary DNA and  
 RT deduced amino acid sequence and localization of its gene (VLDLR) to  
 RT chromosome band 9p24 by fluorescence in situ hybridization.";  
 RL Genomics 20:298-300(1994).  
 RN [5]  
 RP VARIANTS ILE-59 AND LYS-379.  
 RX MEDLINE=93318093; PubMed=10391209;  
 RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,  
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
 RA Lander E.S.;  
 RT "Characterization of single-nucleotide polymorphisms in coding regions  
 RT of human genes.";  
 RL Nat. Genet. 22:231-238(1999).  
 RN [6]  
 RP ERRATUM.  
 RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,  
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
 RA Lander E.S.;  
 RL Nat. Genet. 23:373-373(1999).  
 CC -!- FUNCTION: Binds VLDL and transports it into cells by endocytosis.  
 CC In order to be internalized, the receptor-ligand complexes must  
 CC first cluster into clathrin-coated pits. Binding to Reelin induces  
 CC tyrosine phosphorylation of Dab1 and modulation of Tau  
 CC phosphorylation (By similarity).  
 CC -!- SUBUNIT: Binds to the extracellular matrix protein Reelin (By

```

CC similarity). LOCATION: Type I membrane protein.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoID=P98155-1; Sequence=Displayed;
CC Name=Short;
CC IsoID=P98155-2; Sequence=VSP_004304;
CC -!- TISSUE SPECIFICITY: ABUNDANT IN HEART AND SKELETAL MUSCLE; ALSO
CC OVARY AND KIDNEY; NOT IN LIVER.
CC -!- SIMILARITY: Contains 8 LDL-receptor class A domains.
CC -!- SIMILARITY: Contains 6 LDL-receptor class B domains.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licenses@isb-sib.ch).
CC -----
CC EMBL; L20470; AAA33684.1; -;
CC EMBL; D16532; BAA03969.1; -; JOINED.
CC EMBL; D16495; BAA03969.1; -; JOINED.
CC EMBL; D16508; BAA03969.1; JOINED.
CC EMBL; D16510; BAA03969.1; JOINED.
CC EMBL; D16514; BAA03969.1; JOINED.
CC EMBL; D16516; BAA03969.1; JOINED.
CC EMBL; D16518; BAA03969.1; JOINED.
CC EMBL; D16520; BAA03969.1; JOINED.
CC EMBL; D16522; BAA03969.1; JOINED.
CC EMBL; D16523; BAA03969.1; JOINED.
CC EMBL; D16524; BAA03969.1; JOINED.
CC EMBL; D16525; BAA03969.1; JOINED.
CC EMBL; D16526; BAA03969.1; JOINED.
CC EMBL; D16527; BAA03969.1; JOINED.
CC EMBL; D16528; BAA03969.1; JOINED.
CC EMBL; D16529; BAA03969.1; JOINED.
CC EMBL; D16530; BAA03969.1; JOINED.
CC EMBL; D16531; BAA03969.1; JOINED.
CC EMBL; S73849; AAB31735.1; -;
CC EMBL; D16493; BAA03945.1; -;
CC EMBL; D16494; BAA03946.1; -;
CC EMBL; L22431; AAA61344.1; -;
CC PIR; A49729; A49729.
CC HSSP; P01130; LAJJ.
CC Genew; HGNC:12698; VLDLR.
CC MTM; 192977; -;
CC GO; GO:0005866; C:plasma membrane; TAS.
CC GO; GO:0005041; F:low-density lipoprotein receptor activity; TAS.
CC GO; GO:0007613; P:memory; TAS.
CC GO; GO:0007399; P:neurogenesis; TAS.
CC GO; GO:0007165; P:signal transduction; TAS.
CC InterPro; IPR000152; ASx hydroxyl_S.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR006209; EGF_Like.
CC InterPro; IPR002172; LDL_receptor_A.
CC InterPro; IPR000033; LDL_receptor_rep.
CC Pfam; PF00008; EGF_2.
CC Pfam; PF00057; ldl_recept_a; 8.
CC Pfam; PF00058; ldl_recept_b; 5.
CC PRINTS; PR00261; LDLRECEPTOR.
CC SMART; SM00179; EGF_CA; 2.
CC SMART; SM00192; LDLa; 8.
CC SMART; SM00135; LY; 5.
CC PROSITE; PS00010; ASX_HYDROXYL; 2.
CC PROSITE; PS00022; EGF_1; FALSE_NEG.
CC PROSITE; PS01186; EGF_2; 3.
CC PROSITE; PS00026; EGF_3; 2.
CC PROSITE; PS01187; EGF_CA; 1.
CC PROSITE; PS01209; LDLRA_1; 8.
CC PROSITE; PS00068; LDLRA_2; 8.

```

```

KW Glycoprotein; VLDL; Cholesterol metabolism; lipid transport;
KW Endocytosis; Coated pits; Transmembrane; Receptor; Signal;
KW EGF-like domain; Repeat; Alternative splicing; Polymorphism.
FT SIGNAL 1
FT CHAIN 28 873
FT DOMAIN 28 797
FT TRANSMEM 798 819
FT DOMAIN 820 873
FT DOMAIN 31 69
FT DOMAIN 70 110
FT DOMAIN 111 151
FT DOMAIN 152 190
FT DOMAIN 191 231
FT DOMAIN 237 275
FT DOMAIN 276 314
FT DOMAIN 316 355
FT DOMAIN 356 395
FT DOMAIN 396 435
FT REPEAT 439 480
FT REPEAT 481 524
FT REPEAT 525 567
FT REPEAT 568 611
FT REPEAT 612 654
FT REPEAT 655 696
FT DOMAIN 702 750
FT DOMAIN 751 790
FT SITE 832 837
FT DISULFID 33 45
FT DISULFID 40 58
FT DISULFID 52 67
FT DISULFID 72 84
FT DISULFID 79 97
FT DISULFID 91 108
FT DISULFID 113 127
FT DISULFID 120 140
FT DISULFID 134 149
FT DISULFID 154 166
FT DISULFID 161 179
FT DISULFID 173 188
FT DISULFID 193 205
FT DISULFID 200 218
FT DISULFID 212 229
FT DISULFID 239 251
FT DISULFID 246 264
FT DISULFID 258 273
FT DISULFID 278 290
FT DISULFID 285 303
FT DISULFID 297 312
FT DISULFID 318 331
FT DISULFID 326 344
FT DISULFID 338 355
FT DISULFID 360 371
FT DISULFID 367 380
FT DISULFID 382 394
FT DISULFID 400 410
FT DISULFID 406 419
FT DISULFID 421 434

```

```

Query Match 17.0%; Score 82.5; DB 1; Length 873;
Best Local Similarity 31.7%; Pred. No. 1.1;
Matches 26; Conservative 7; Mismatches 38; Indels 11; Gaps 4;

QY 5 OCGENEKYDS-----CGSKEC-DKKCKYDGVVEEDERPNVPCLVVCHQD-CVCEE 54
DB 228 QCGROPVHTKCPASEIQCGSGECHKWRCGDFCDCKDGSDEVNCPRTCEPTQFECED 287
QY 55 GFYRNKDKCVSABDCELDND 76
DB 288 GSCIHGSRQCNIGRDC-VDGSD 308

```

Search completed: September 24, 2004, 07:34:56

Fri Sep 24 08:54:02 2004

us-09-498-556c-59.sep04.rsp

Page 16

Job time : 31.1789 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 07:28:32 ; Search time 16.9474 Seconds  
(without alignments)  
116.704 Million cell updates/sec

Title: US-09-498-556C-79

Perfect score: 32  
Sequence: 1 XXGFYRN 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1590s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	93.8	5	2 AAR91721	Aar91721 NAP subse
2	30	93.8	5	2 AAY30440	Aay30440 Nematode
3	30	93.8	5	3 AAB15325	Aab15325 NAP domai
4	30	93.8	7	2 AAR91730	Aar91730 NAP subse
5	30	93.8	7	2 AAY30449	Aay30449 Nematode
6	30	93.8	7	3 AAB15334	Aab15334 NAP domai
7	30	93.8	78	2 AAY30420	Aay30420 Mature ne
8	30	93.8	78	3 AAB15305	Aab15305 A. caninu
9	30	93.8	82	2 AAY30422	Aay30422 Mature ne
10	30	93.8	82	2 AAY30399	Aay30399 Nematode
11	30	93.8	82	3 AAB15293	Aab15293 A. ceylan
12	30	93.8	82	3 AAB15307	Aab15307 A. ceylan
13	30	93.8	83	2 AAY30425	Aay30425 Mature ne
14	30	93.8	84	2 AAY30432	Aay30432 Mature ne
15	30	93.8	84	2 AAY30423	Aay30423 Mature ne
16	30	93.8	84	2 AAY30424	Aay30424 Mature ne
17	30	93.8	84	3 AAB15317	Aab15317 A. caninu
18	30	93.8	86	4 AAW91998	Aaw91998 Human dig
19	30	93.8	89	2 AAY30421	Aay30421 Mature ne
20	30	93.8	91	2 AAR91701	Aar91701 AcanaPc2.
21	30	93.8	91	2 AAY30393	Aay30393 Nematode
22	30	93.8	91	2 AAY30454	Aay30454 Nematode
23	30	93.8	91	3 AAB15346	Aab15346 A. caninu
24	30	93.8	97	2 AAR91709	Aar91709 AcanaP31.
25	30	93.8	97	2 AAY30407	Aay30407 Nematode

26	30	93.8	97	4 ABB14986	Abb14986 Human ner
27	30	93.8	102	2 AAY30401	Aay30401 Nematode
28	30	93.8	102	3 AAB15295	Aab15295 A. duoden
29	30	93.8	108	2 AAY30398	Aay30398 Nematode
30	30	93.8	108	3 AAB15292	Aab15292 A. ceylan
31	30	93.8	124	3 AAY70530	Aay70530 Maize pla
32	30	93.8	138	4 ABB03987	Abb03987 Human mus
33	30	93.8	138	6 ABB13281	Abb13281 Novel hum
34	30	93.8	161	2 AAY30438	Aay30438 Mature ne
35	30	93.8	162	2 AAY30436	Aay30436 Mature ne
36	30	93.8	162	2 AAY30437	Aay30437 Mature ne
37	30	93.8	162	3 AAB15321	Aab15321 A. caninu
38	30	93.8	162	3 AAB15322	Aab15322 A. caninu
39	30	93.8	171	2 AAY30435	Aay30435 Mature ne
40	30	93.8	180	2 AAR91713	Aar91713 AduNAP7.
41	30	93.8	180	2 AAY30388	Aay30388 Nematode
42	30	93.8	181	2 AAR91712	Aar91712 AcanaP47.
43	30	93.8	181	2 AAR91711	Aar91711 AcanaP45.
44	30	93.8	181	2 AAY30409	Aay30409 Nematode
45	30	93.8	181	2 AAY30410	Aay30410 Nematode

## ALIGNMENTS

RESULT 1  
AAR91721  
ID AAR91721 standard; protein; 5 AA.  
XX

AC AAR91721;  
DT 25-MAR-2003 (revised)  
DT 17-NOV-1996 (first entry)  
XX  
DE NAP subsequence.  
XX  
KW AcANAP; HPOANP; NamNAP; AcenAP; AduNAP; anticoagulant;  
KW nematode-extracted anticoagulant protein; serine protease; nematode;  
KW thrombosis; parasitic worm.  
XX  
OS Synthetic.  
XX  
PN WO9612021-A2.  
XX  
PD 25-APR-1996.  
XX  
PF 17-OCT-1995; 95WO-US013231.  
XX

PR 18-OCT-1994; 94US-00326110.  
PR 05-JUN-1995; 95US-00461965.  
PR 05-JUN-1995; 95US-00465380.  
PR 05-JUN-1995; 95US-00486397.  
PR 05-JUN-1995; 95US-00486399.  
XX  
(CORV-) CORVAS INT INC.  
XX  
Vlasuk GP, Stanssens PEH, Messens JHL, Lauwereys MJ, Laroche YR;  
Jespers LS, Ganssems YGJ, Moyle M, Bergum PW;  
XX  
WPI, 1996-222007/22.  
XX  
Proteins with anticoagulant and/or serine protease inhibitory activity -  
isolated from nematodes and useful to inhibit blood coagulation.  
XX  
Claim 10; Page 144; 243pp; English.  
XX  
Proteins with anticoagulant and/or serine protease inhibitory activity,  
isolated from nematodes, are useful to inhibit blood coagulation. The  
proteins can be added to blood collection tubes defining the collection  
of mammalian plasma. They are also useful to prevent or inhibit  
thrombosis, and may be given alone or in combination with other  
therapeutic or in vivo diagnostic agents. The proteins can serve as  
immunogens to raise antibodies for use in the diagnosis and

CC identification of NAP concn. levels in biological fluids, e.g. to detect  
 CC mammalian infection with a parasitic worm. They can also be used as  
 CC immunogens in prophylactic and therapeutic vaccines against parasitic  
 CC worm infection. The proteins may double the clotting time of human plasma  
 CC in prothrombin time assays when present at 10-50 nMol, and double the  
 CC clotting time of human plasma in activated partial thrombin time assays  
 CC when present at 10-100 nMol. The anticoagulant proteins are pref. derived  
 CC from *Ancylostoma caninum*, *A. ceylanicum*, *A. duodenale*, *Necator americanus*  
 CC or *Heligmosomoides polygyrus*. The proteins pref. have 2 NAP domains and  
 CC specifically inhibit the catalytic activity of the factor VIIa/TF complex  
 CC in the presence of factor Xa or a catalytically inactive factor Xa  
 CC deriv., do not specifically inhibit the activation of factor VIIa in the  
 CC absence of TF and do not specifically inhibit prothrombinase. Proteins  
 CC given in AAR91720-R91732 are preferred subsequences of a generic NAP  
 CC sequence. (Updated on 25-MAR-2003 to correct PI field.)  
 XX

SQ Sequence 5 AA;  
 Query Match 93.8%; Score 30; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
 |||||  
 Db 1 GFYRN 5

RESULT 2  
 AAY30440  
 ID AAY30440 standard; peptide; 5 AA.

XX AC AAY30440;  
 XX DT 15-NOV-1999 (first entry)  
 XX DE Nematode extracted anticoagulant protein fragment.  
 XX DE Nematode extracted anticoagulant protein; NAP; anticoagulant;  
 XX KW serine protease inhibitor; NAP domain; factor VIIa/TF.  
 XX KW Unidentified.  
 XX OS US955294-A.  
 XX PN 21-SEP-1999.  
 XX PF 19-APR-1996; 96US-00634641.  
 XX PR 18-OCT-1994; 94US-00326110.  
 XX PR 05-JUN-1995; 95US-00461965.  
 XX PR 05-JUN-1995; 95US-00465380.  
 XX PR 05-JUN-1995; 95US-00486397.  
 XX PR 05-JUN-1995; 95US-00486399.  
 XX PR 17-OCT-1995; 95WO-US013231.  
 XX PA (CORV-) CORVAS INT INC.

XX PI Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;  
 XX PI Bergum PW, Messens JHL, Laroche YR, Vlasuk GP;  
 XX WI; 1999-539569/45.

XX PT Screening an isolated protein for Nematode-extracted Anticoagulant  
 XX PT Protein domains.

XX PS Disclosure; Col 151; 197pp; English.

XX CC The present sequence represents a nematode extracted anticoagulant  
 CC protein (NAP) fragment. The protein has activity as an anticoagulant  
 CC and/or serine protease inhibitor. The protein contains at least one NAP  
 CC domain which has selective inhibitory activity for factor VIIa/TF. The  
 CC specification describes a method for screening an isolated protein at  
 CC least one domain for factor VIIa/TF selective inhibitory activity. The

CC method comprises determining the time to clotting effected by a  
 CC concentration of the isolated protein in an ex vivo prothrombin time (PT)  
 CC assay and an ex vivo activated partial thromboplastin time (APTT) assay;  
 CC calculating prolongation of clotting effected by the isolated protein in  
 CC each of the PT and APTT assay, with respect to a baseline clotting value  
 CC for each assay, where prolongation of clotting is calculated as fold  
 CC elevation of clotting time relative to a baseline clotting value, where a  
 CC doubling of clotting time is deemed a two-fold elevation; and calculating  
 CC a PT to APTT prolongation ratio, where a ratio at least one is indicative  
 CC of factor VIIa/TF inhibitory activity. The method is useful for  
 CC determining if a protein has factor VIIa/TF inhibitory activity  
 XX

SQ Sequence 5 AA;

Query Match 93.8%; Score 30; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
 |||||  
 Db 1 GFYRN 5

RESULT 3  
 AAB15325  
 ID AAB15325 standard; peptide; 5 AA.

XX AC AAB15325;  
 XX DT 19-DEC-2000 (first entry)  
 XX DE NAP domain fragment #2.  
 XX KW Nematode-extracted anticoagulant protein; NAP domain; blood clotting;  
 XX KW canine hookworm; thrombosis; vaccine.  
 XX OS Unidentified.  
 XX PN US6087487-A.  
 XX PD 11-JUL-2000.  
 XX PF 12-FEB-1999; 99US-00249451.  
 XX PR 18-OCT-1994; 94US-00326110.  
 XX PR 05-JUN-1995; 95US-00461965.  
 XX PR 05-JUN-1995; 95US-00465380.  
 XX PR 05-JUN-1995; 95US-00486397.  
 XX PR 05-JUN-1995; 95US-00486399.  
 XX PR 17-OCT-1995; 95WO-US013231.  
 XX PR 17-APR-1997; 97US-00809455.

XX PA (CORV-) CORVAS INT INC.

XX PI Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;  
 XX PI Bergum PW, Messens JHL, Laroche YR, Vlasuk GP;  
 XX WI; 2000-531359/48.

XX PT New cDNA molecule encoding a protein having factor Xa inhibitory activity  
 XX PT for preventing and treating blood clotting disorders, comprises nematode-  
 XX PT extracted anticoagulant protein domains.

XX PS Claim 3; Col 285; 197pp; English.

XX CC The present sequence is a fragment of the NAP domain (see AAB15347),  
 XX CC which is found in all nematode-extracted anticoagulant proteins (NAPe).  
 XX CC Proteins of this kind have been shown to be effective at preventing blood  
 XX CC clotting without causing excessive bleeding. They can be used in blood  
 XX CC collection tubes to aid the isolation of plasma from the blood, to  
 XX CC prevent thrombosis which may be linked to the rupture of an  
 XX CC atherosclerotic plaque, acute myocardial infarction, angina, thrombolytic  
 XX CC therapy, percutaneous transluminal coronary angioplasty, disseminated

CC intravascular coagulopathy, infection, cancer and septic shock, and to  
 CC produce antibodies. In the latter instance, the antibodies can be raised  
 CC in order to detect infection by nematodes (the coding sequences can also  
 CC be used for this) or as diagnostic tests. The proteins can also be used  
 CC as vaccines against nematode parasites  
 XX  
 SQ Sequence 5 AA;  
 Query Match 93.8%; Score 30; DB 3; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 GFYRN 7  
 DB 1 GFYRN 5  
 RESULT 4  
 AAR91730  
 ID AAR91730 standard; protein; 7 AA.  
 AC AAR91730;  
 XX  
 XX 25-MAR-2003 (revised)  
 DT 17-NOV-1996 (first entry)  
 DT  
 XX  
 DE NAP subsequence.  
 XX  
 XX AcaNAP; HpocNAP; NamNAP; AcaNAP; AduNAP; anticoagulant;  
 XX nematode-extracted anticoagulant protein; serine protease; nematode;  
 KW thrombosis; parasitic worm.  
 XX  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 1 /label= Glu, Asp, OTHER  
 FT /note= "at least one of residue 1 or residue 2 is Glu or  
 FT Asp"  
 FT Misc-difference 2 /label= Glu, Asp, OTHER  
 FT /note= "at least one of residue 1 or residue 2 is Glu or  
 FT Asp"  
 XX  
 XX WO9612021-A2.  
 XX  
 XX 25-APR-1996.  
 XX  
 XX 17-OCT-1995; 95WO-US013231.  
 XX  
 XX 18-OCT-1994; 94US-00326110.  
 XX 05-JUN-1995; 95US-00461965.  
 XX 05-JUN-1995; 95US-00465380.  
 XX 05-JUN-1995; 95US-00486397.  
 XX 05-JUN-1995; 95US-00486399.  
 XX  
 XX (CORV-) CORVAS INT INC.  
 XX  
 XX Vlausk GP, Stanssens PEH, Messens JHL, Lauwereys MJ, Laroche YR;  
 XX Jespers LS, Gansemans YGJ, Moyle M, Bergum PW;  
 XX WPI; 1996-222007/22.  
 XX  
 XX Proteins with anticoagulant and/or serine protease inhibitory activity -  
 XX isolated from nematodes and useful to inhibit blood coagulation.  
 XX  
 XX Claim 26; Page 147; 243pp; English.  
 XX  
 XX Proteins with anticoagulant and/or serine protease inhibitory activity,  
 CC isolated from nematodes, are useful to inhibit blood coagulation. The  
 CC proteins can be added to blood collection tubes defining the collection  
 CC of mammalian plasma. They are also useful to prevent or inhibit  
 CC thrombosis, and may be given alone or in combination with other

CC therapeutic or in vivo diagnostic agents. The proteins can serve as  
 CC immunogens to raise antibodies for use in the diagnosis and  
 CC identification of NAP concn. levels in biological fluids, e.g. to detect  
 CC mammalian infection with a parasitic worm. They can also be used as  
 CC immunogens in prophylactic and therapeutic vaccines against parasitic  
 CC worm infection. The proteins may double the clotting time of human plasma  
 CC in prothrombin time assays when present at 10-50 nMol, and double the  
 CC clotting time of human plasma in activated partial thrombin time assays  
 CC when present at 10-100 nMol. The anticoagulant proteins are pref. derived  
 CC from *Ancylostoma caninum*, *A. ceylanicum*, *A. duodenale*, *Necator americanus*  
 CC or *Heligmosomoides polygyrus*. The proteins pref. have 2 NAP domains and  
 CC specifically inhibit the catalytic activity of the factor VIIa/TF complex  
 CC in the presence of factor Xa or a catalytically inactive factor Xa  
 CC deriv., do not specifically inhibit the activation of factor VIIa in the  
 CC absence of TF and do not specifically inhibit prothrombinase. Proteins  
 CC given in AAR91720-R91732 are preferred subsequences of a generic NAP  
 CC sequence. (Updated on 25-MAR-2003 to correct PI field.)  
 XX  
 SQ Sequence 7 AA;

Query Match 93.8%; Score 30; DB 2; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
 DB 3 GFYRN 7

RESULT 5  
 AAY30449  
 ID AAY30449 standard; peptide; 7 AA.  
 AC AAY30449;  
 XX  
 XX 15-NOV-1999 (first entry)  
 DT  
 XX Nematode extracted anticoagulant protein fragment.  
 DE  
 XX Nematode extracted anticoagulant protein; NAP; anticoagulant;  
 KW serine protease inhibitor; NAP domain; factor VIIa/TF.  
 XX  
 OS Unidentified.  
 XX  
 XX US5955294-A.  
 XX  
 XX 21-SEP-1999.  
 XX  
 XX 19-APR-1996; 96US-00634641.  
 XX  
 XX 18-OCT-1994; 94US-00326110.  
 XX 05-JUN-1995; 95US-00461965.  
 XX 05-JUN-1995; 95US-00485380.  
 XX 05-JUN-1995; 95US-00486397.  
 XX 05-JUN-1995; 95US-00486399.  
 XX 17-OCT-1995; 95WO-US013231.  
 XX  
 XX (CORV-) CORVAS INT INC.

XX Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;  
 XX Bergum PW, Messens JHL, Laroche YR, Vlausk GP;  
 XX WPI; 1999-539569/45.

XX Screening an isolated protein for Nematode-extracted Anticoagulant  
 XX Protein domains.

XX Disclosure; Col 155; 197pp; English.

XX The present sequence represents a nematode extracted anticoagulant  
 CC protein (NAP) fragment. The protein has activity as an anticoagulant  
 CC and/or serine protease inhibitor. The protein contains at least one NAP  
 CC domain which has selective inhibitory activity for factor VIIa/TF. The

CC specification describes a method for screening an isolated protein at  
 CC least one domain for factor VIIa/TF selective inhibitory activity. The  
 CC method comprises determining the time to clotting effected by a  
 CC concentration of the isolated protein in an ex vivo prothrombin time (PT)  
 CC assay and an ex vivo activated partial thromboplastin time (APTT) assay;  
 CC calculating prolongation of clotting effected by the isolated protein in  
 CC each of the PT and APTT assay, with respect to a baseline clotting value  
 CC for each assay, where prolongation of clotting is calculated as fold  
 CC elevation of clotting time relative to a baseline clotting value, where a  
 CC doubling of clotting time is deemed a two-fold elevation; and calculating  
 CC a PT to APTT prolongation ratio, where a ratio at least one is indicative  
 CC of factor VIIa/TF inhibitory activity. The method is useful for  
 CC determining if a protein has factor VIIa/TF inhibitory activity  
 XX  
 XX

SQ Sequence 7 AA;

Query Match 93.8%; Score 30; DB 2; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
 |||||  
 Db 3 GFYRN 7

RESULT 6  
 AAB15334  
 ID AAB15334 standard; peptide; 7 AA.

XX AC AAB15334;

XX DT 19-DEC-2000 (first entry)

XX DE NAP domain fragment #11.

XX KW Nematode-extracted anticoagulant protein; NAP domain, blood clotting;  
 XX KW canine hookworm; thrombosis; vaccine.

XX OS Unidentified.

XX FH Key Location/Qualifiers

XX FT Misc-difference 1..2

XX FT /label= OTHER

XX FT /note= "one of these two residues must be either Glu or

XX FT Asp"

XX US087487-A.

XX 11-JUL-2000.

XX 12-FEB-1999; 99US-00249451.

XX 18-OCT-1994; 94US-00326110.

XX 05-JUN-1995; 95US-00461965.

XX 05-JUN-1995; 95US-00465380.

XX 05-JUN-1995; 95US-00486397.

XX 05-JUN-1995; 95US-00486399.

XX 17-OCT-1995; 95WO-US013231.

XX 17-APR-1997; 97US-00809455.

XX (CORV-) CORVAS INT INC.

XX Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;

XX Bergum PW, Messens JHL, Laroche YR, Vlasuk GP;

XX WPI; 2000-531359/48.

XX New cDNA molecule encoding a protein having factor Xa inhibitory activity  
 XX for preventing and treating blood clotting disorders, comprises nematode-  
 XX extracted anticoagulant protein domains.

XX Claim 4; Col 286; 197pp; English.

CC The present sequence is a fragment of the NAP domain (see AAB15347),  
 CC which is found in all nematode-extracted anticoagulant proteins (NAPs).  
 CC proteins of this kind have been shown to be effective at preventing blood  
 CC clotting without causing excessive bleeding. They can be used in blood  
 CC collection tubes to aid the isolation of plasma from the blood, to  
 CC prevent thrombosis which may be linked to the rupture of an  
 CC atherosclerotic plaque, acute myocardial infarction, angina, thrombolytic  
 CC therapy, percutaneous transluminal coronary angioplasty, disseminated  
 CC intravascular coagulopathy, infection, cancer and septic shock, and to  
 CC produce antibodies. In the latter instances, the antibodies can be raised  
 CC in order to detect infection by nematodes (the coding sequences can also  
 CC be used for this) or as diagnostic tests. The proteins can also be used  
 CC as vaccines against nematode parasites  
 XX  
 XX

SQ Sequence 7 AA;

Query Match 93.8%; Score 30; DB 3; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
 |||||  
 Db 3 GFYRN 7

RESULT 7

AAAY30420

ID AAAY30420 standard; protein; 78 AA.

XX AC AAAY30420;

XX DT 15-NOV-1999 (first entry)

XX DE Mature nematode extracted anticoagulant protein AcanAP31.

XX KW Nematode extracted anticoagulant protein; NAP; anticoagulant;  
 XX KW serine protease inhibitor; NAP domain; factor VIIa/TF.

XX OS Ancylostoma caninum.

XX PN US5955294-A.

XX PD 21-SEP-1999.

XX PF 19-APR-1996; 96US-00634641.

XX PR 18-OCT-1994; 94US-00326110.

XX PR 05-JUN-1995; 95US-00461965.

XX PR 05-JUN-1995; 95US-00465380.

XX PR 05-JUN-1995; 95US-00486397.

XX PR 05-JUN-1995; 95US-00486399.

XX PR 17-OCT-1995; 95WO-US013231.

XX (CORV-) CORVAS INT INC.

XX Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;

XX Bergum PW, Messens JHL, Laroche YR, Vlasuk GP;

XX WPI; 1999-539569/45.

XX Screening an isolated protein for Nematode-extracted Anticoagulant

XX Protein domains.

XX Disclosure; Col 133-134; 197pp; English.

XX The present sequence represents a nematode extracted anticoagulant  
 CC protein (NAP). The protein has activity as an anticoagulant and/or serine  
 CC protease inhibitor. The protein contains at least one NAP domain which  
 CC has selective inhibitory activity for factor VIIa/TF. The specification  
 CC describes a method for screening an isolated protein at least one domain  
 CC for factor VIIa/TF selective inhibitory activity. The method comprises  
 CC determining the time to clotting effected by a concentration of the  
 CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo

CC activated partial thromboplastin time (aPTT) assay; calculating  
 CC prolongation of clotting effected by the isolated protein in each of the  
 CC PT and aPTT assay, with respect to a baseline clotting value for each  
 CC assay, where prolongation of clotting is calculated as fold elevation of  
 CC clotting time relative to a baseline clotting value, where a doubling of  
 CC clotting time is deemed a two-fold elevation; and calculating a PT to  
 CC aPTT prolongation ratio, where a ratio at least one is indicative of  
 CC factor VIIa/TF inhibitory activity. The method is useful for determining  
 CC if a protein has factor VIIa/TF inhibitory activity  
 XX  
 SQ Sequence 78 AA;

Query Match 93.8%; Score 30; DB 2; Length 78;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
 |||||  
 Db 54 GFYRN 58

RESULT 8  
 AAB15305  
 ID AAB15305 standard; protein; 78 AA.  
 XX AC  
 XX AAB15305;  
 AC  
 DT 06-AUG-2003 (revised)  
 DT 19-DEC-2000 (first entry)  
 XX  
 XX A. caninum nematode-extracted anticoagulant protein AcanAP31,42,46.  
 DE  
 XX Nematode-extracted anticoagulant protein; AcanAP31; AcanAP42; AcanAP46;  
 KW canine hookworm; Blood clotting; thrombosis; vaccine.  
 KW  
 XX Ancylostoma caninum.  
 OS  
 XX US6087487-A.  
 PN  
 XX 11-JUL-2000.  
 XX  
 XX 12-FEB-1999; 99US-00249451.  
 XX  
 PR 18-OCT-1994; 94US-00326110.  
 PR 05-JUN-1995; 95US-00461965.  
 PR 05-JUN-1995; 95US-00465380.  
 PR 05-JUN-1995; 95US-00486397.  
 PR 05-JUN-1995; 95US-00486399.  
 PR 17-OCT-1995; 95WO-US013231.  
 PR 17-APR-1997; 97US-00809455.  
 XX  
 XX (CORV-) CORVAS INT INC.

PI Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;  
 PI Bergum PW, Messens JHL, Laroche YR, Vlasuk GP;  
 XX  
 XX WPI; 2000-531359/48.  
 DR N-PSDB; AAA73378.  
 XX  
 XX New cDNA molecule encoding a protein having factor Xa inhibitory activity  
 PT for preventing and treating blood clotting disorders, comprises nematode-  
 PT extracted anticoagulant protein domains.  
 XX  
 XX Disclosure; Fig 16; 197pp; English.  
 XX  
 CC The present sequence comprises the Ancylostoma caninum nematode-  
 CC extracted anticoagulant proteins AcanAP31, AcanAP42 and AcanAP46.  
 CC Proteins of this kind have been shown to be effective at preventing blood  
 CC clotting without causing excessive bleeding. The proteins can be used in  
 CC blood collection tubes to aid the isolation of plasma from the blood, to  
 CC prevent thrombosis which may be linked to the rupture of an  
 CC atherosclerotic plaque, acute myocardial infarction, angina, thrombolytic  
 CC therapy, percutaneous transluminal coronary angioplasty, disseminated

CC intravascular coagulopathy, infection, cancer and septic shock, and to  
 CC produce antibodies. In the latter instance, the antibodies can be raised  
 CC in order to detect infection by nematodes (the coding sequence can also  
 CC be used for this) or as diagnostic tests. The proteins can also be used  
 CC as a vaccine against nematode parasites. (Updated on 06-AUG-2003 to  
 CC correct OS field.)  
 XX  
 SQ Sequence 78 AA;

Query Match 93.8%; Score 30; DB 3; Length 78;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
 |||||  
 Db 54 GFYRN 58

RESULT 9  
 AAY30422  
 ID AAY30422 standard; protein; 82 AA.  
 XX AC  
 XX AAY30422;  
 AC  
 DT 15-NOV-1999 (first entry)  
 XX  
 XX Mature nematode extracted anticoagulant protein AcanAP4d2.  
 DE  
 XX Nematode extracted anticoagulant protein; NAP; anticoagulant;  
 KW serine protease inhibitor; NAP domain; factor VIIa/TF.  
 KW  
 XX Ancylostoma ceylanicum.  
 OS  
 XX US955294-A.  
 PN  
 XX 21-SEP-1999.  
 XX  
 XX 19-APR-1996; 96US-00634641.  
 XX  
 PR 18-OCT-1994; 94US-00326110.  
 PR 05-JUN-1995; 95US-00461965.  
 PR 05-JUN-1995; 95US-00465380.  
 PR 05-JUN-1995; 95US-00486397.  
 PR 05-JUN-1995; 95US-00486399.  
 PR 17-OCT-1995; 95WO-US013231.  
 XX  
 XX (CORV-) CORVAS INT INC.

PI Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;  
 PI Bergum PW, Messens JHL, Laroche YR, Vlasuk GP;  
 XX  
 XX WPI; 1999-539569/45.  
 XX  
 XX Screening an isolated protein for Nematode-extracted Anticoagulant  
 PT Protein domains.  
 XX  
 XX Disclosure; Col 135-136; 197pp; English.  
 XX  
 CC The present sequence represents a nematode extracted anticoagulant  
 CC protein (NAP). The protein has activity as an anticoagulant and/or serine  
 CC protease inhibitor. The protein contains at least one NAP domain which  
 CC has selective inhibitory activity for factor VIIa/TF. The specification  
 CC describes a method for screening an isolated protein at least one domain  
 CC for factor VIIa/TF selective inhibitory activity. The method comprises  
 CC determining the time to clotting effected by a concentration of the  
 CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo  
 CC activated partial thromboplastin time (aPTT) assay; calculating  
 CC prolongation of clotting effected by the isolated protein in each of the  
 CC PT and aPTT assay, with respect to a baseline clotting value for each  
 CC assay, where prolongation of clotting is calculated as fold elevation of  
 CC clotting time relative to a baseline clotting value, where a doubling of  
 CC clotting time is deemed a two-fold elevation; and calculating a PT to  
 CC aPTT prolongation ratio, where a ratio at least one is indicative of

CC factor VIIa/TF inhibitory activity. The method is useful for determining  
 CC if a protein has factor VIIa/TF inhibitory activity

XX Query Match 93.8%; Score 30; DB 2; Length 82;  
 XX Best Local Similarity 100.0%; Pred. No. 79;  
 SQ Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 SQ Sequence 82 AA;

Query Match 93.8%; Score 30; DB 2; Length 82;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 GFYRN 7  
 DB 51 GFYRN 55

RESULT 10  
 AAY30399  
 ID AAY30399 standard; protein; 82 AA.

XX AC AAY30399;  
 XX DT 15-NOV-1999 (first entry)  
 XX DE Nematode extracted anticoagulant protein AceNAP4d2.  
 XX KW Nematode extracted anticoagulant protein; NAP; anticoagulant;  
 XX KW serine protease inhibitor; NAP domain; factor VIIa/TF.  
 XX OS Ancylostoma ceylanicum.

XX PN US955294-A.  
 XX PD 21-SEP-1999.  
 XX PF 19-APR-1996; 96US-00634641.  
 XX PR 18-OCT-1994; 94US-00326110.  
 XX PR 05-JUN-1995; 95US-00461965.  
 XX PR 05-JUN-1995; 95US-00465380.  
 XX PR 05-JUN-1995; 95US-00486397.  
 XX PR 05-JUN-1995; 95US-00486399.  
 XX PR 17-OCT-1995; 95WO-US013231.  
 XX PA (CORV-) CORVAS INT INC.

XX PI Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;  
 XX PI Bergum PW, Messens JHL, Laroche YR, Vlasuk GP;  
 XX WPI; 1999-539569/45.

XX Screening an isolated protein for Nematode-extracted Anticoagulant  
 XX Protein domains.

XX Disclosure: Col 111-112; 197pp; English.

XX The present sequence represents a nematode extracted anticoagulant  
 XX protein (NAP). The protein has activity as an anticoagulant and/or serine  
 XX protease inhibitor. The protein contains at least one NAP domain which  
 XX has selective inhibitory activity for factor VIIa/TF. The specification  
 XX describes a method for screening an isolated protein at least one domain  
 XX for factor VIIa/TF selective inhibitory activity. The method comprises  
 XX determining the time to clotting effected by a concentration of the  
 XX isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo  
 XX activated partial thromboplastin time (APTT) assay; calculating  
 XX prolongation of clotting effected by the isolated protein in each of the  
 XX PT and APTT assay, with respect to a baseline clotting value for each of  
 XX assay, where prolongation of clotting is calculated as fold elevation of  
 XX clotting time relative to a baseline clotting value, where a doubling of  
 XX clotting time is deemed a two-fold elevation; and calculating a PT to  
 XX factor VIIa/TF inhibitory activity. The method is useful for determining  
 XX if a protein has factor VIIa/TF inhibitory activity

XX Sequence 82 AA;

Query Match 93.8%; Score 30; DB 2; Length 82;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 GFYRN 7  
 DB 51 GFYRN 55

RESULT 11  
 AAB15293  
 ID AAB15293 standard; protein; 82 AA.

XX AC AAB15293;  
 XX DT 06-AUG-2003 (revised)  
 XX DT 19-DEC-2000 (first entry)  
 XX DE A. ceylanicum nematode-extracted anticoagulant protein AceNAP4d2 #1.  
 XX KW Nematode-extracted anticoagulant protein; AceNAP4d2; blood clotting;  
 XX KW canine hookworm; thrombosis; vaccine.

XX OS Ancylostoma ceylanicum.

XX PN US6087487-A.  
 XX PD 11-JUL-2000.  
 XX PF 12-FEB-1999; 99US-00249451.  
 XX PR 18-OCT-1994; 94US-00326110.  
 XX PR 05-JUN-1995; 95US-00461965.  
 XX PR 05-JUN-1995; 95US-00465380.  
 XX PR 05-JUN-1995; 95US-00486397.  
 XX PR 05-JUN-1995; 95US-00486399.  
 XX PR 17-OCT-1995; 95WO-US013231.  
 XX PR 17-APR-1997; 97US-00809455.  
 XX PA (CORV-) CORVAS INT INC.

XX PI Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;  
 XX PI Bergum PW, Messens JHL, Laroche YR, Vlasuk GP;  
 XX WPI; 2000-531359/48.

XX New cDNA molecule encoding a protein having factor Xa inhibitory activity  
 XX for preventing and treating blood clotting disorders, comprises nematode-  
 XX extracted anticoagulant protein domains.

XX Disclosure: Fig 11; 197pp; English.

XX The present sequence is the Ancylostoma ceylanicum nematode-extracted  
 XX anticoagulant protein AceNAP4d2. Proteins of this kind have been shown to  
 XX be effective at preventing blood clotting without causing excessive  
 XX bleeding. The protein can be used in blood collection tubes to aid the  
 XX isolation of plasma from the blood, to prevent thrombosis which may be  
 XX linked to the rupture of an atherosclerotic plaque, acute myocardial  
 XX infarction, angina, thrombolytic therapy, percutaneous transluminal  
 XX coronary angioplasty, disseminated intravascular coagulopathy, infection,  
 XX cancer and septic shock, and to produce antibodies. In the latter  
 XX instance, the antibodies can be raised in order to detect infection by  
 XX nematodes or as diagnostic tests. The protein can also be used as a  
 XX vaccine against nematode parasites. (Updated on 06-AUG-2003 to correct OS  
 XX field.)

XX Sequence 82 AA;

Query Match 93.8%; Score 30; DB 3; Length 82;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



AC AAY30432;  
 XX  
 DT 15-NOV-1999 (first entry)  
 XX DE  
 XX DE Mature nematode extracted anticoagulant protein AcanAPc2.  
 XX KW Nematode extracted anticoagulant protein; NAP, anticoagulant;  
 KW KW serine protease inhibitor; NAP domain; factor VIIa/TF.  
 XX OS Ancylostoma caninum.  
 XX PN US955294-A.  
 XX PD 21-SEP-1999.  
 XX PF 19-APR-1996; 96US-00634641.  
 XX PR 18-OCT-1994; 94US-00326110.  
 XX PR 05-JUN-1995; 95US-00461965.  
 XX PR 05-JUN-1995; 95US-00465380.  
 XX PR 05-JUN-1995; 95US-00486397.  
 XX PR 05-JUN-1995; 95US-00486399.  
 XX PR 17-OCT-1995; 95WO-US013231.  
 XX PA (CORV-) CORVAS INT INC.  
 XX PI Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;  
 PI Bergum PW, Messens JHL, Laroche YR, Vlasuk GP;  
 XX WPI; 1999-539569/45.  
 XX Screening an isolated protein for Nematode-extracted Anticoagulant  
 PT Protein domains.  
 XX Disclosure; Col 142-144; 197pp; English.  
 XX The present sequence represents a nematode extracted anticoagulant  
 CC protein (NAP). The protein has activity as an anticoagulant and/or serine  
 CC protease inhibitor. The protein contains at least one NAP domain which  
 CC has selective inhibitory activity for factor VIIa/TF. The specification  
 CC describes a method for screening an isolated protein at least one domain  
 CC for factor VIIa/TF selective inhibitory activity. The method comprises  
 CC determining the time to clotting effected by a concentration of the  
 CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo  
 CC activated partial thromboplastin time (APTT) assay; calculating  
 CC prolongation of clotting effected by the isolated protein in each of the  
 CC PT and APTT assay, with respect to a baseline clotting value for each  
 CC assay, where prolongation of clotting is calculated as fold elevation of  
 CC clotting time relative to a baseline clotting value, where a doubling of  
 CC clotting time is deemed a two-fold elevation; and calculating a PT to  
 CC APTT prolongation ratio, where a ratio at least one is indicative of  
 CC factor VIIa/TF inhibitory activity. The method is useful for determining  
 CC if a protein has factor VIIa/TF inhibitory activity  
 XX Sequence 84 AA;  
 Query Match 93.8%; Score 30; DB 2; Length 84;  
 Best Local Similarity 100.0%; Pred. No. 81;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 GFYRN 7  
 Db 55 GFYRN 59  
 RESULT 15  
 ID AAY30423  
 XX AAY30423 standard; protein; 84 AA.  
 XX AAY30423;  
 XX AAY30423;  
 DT 15-NOV-1999 (first entry)  
 XX

DE Mature nematode extracted anticoagulant protein AcanAP45d1.  
 XX  
 KW Nematode extracted anticoagulant protein; NAP, anticoagulant;  
 KW KW serine protease inhibitor; NAP domain; factor VIIa/TF.  
 XX OS Ancylostoma caninum.  
 XX PN US955294-A.  
 XX PD 21-SEP-1999.  
 XX PF 19-APR-1996; 96US-00634641.  
 XX PR 18-OCT-1994; 94US-00326110.  
 XX PR 05-JUN-1995; 95US-00461965.  
 XX PR 05-JUN-1995; 95US-00465380.  
 XX PR 05-JUN-1995; 95US-00486397.  
 XX PR 05-JUN-1995; 95US-00486399.  
 XX PR 17-OCT-1995; 95WO-US013231.  
 XX PA (CORV-) CORVAS INT INC.  
 XX PI Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;  
 PI Bergum PW, Messens JHL, Laroche YR, Vlasuk GP;  
 XX WPI; 1999-539569/45.  
 XX Screening an isolated protein for Nematode-extracted Anticoagulant  
 PT Protein domains.  
 XX Disclosure; Col 135-136; 197pp; English.  
 XX The present sequence represents a nematode extracted anticoagulant  
 CC protein (NAP). The protein has activity as an anticoagulant and/or serine  
 CC protease inhibitor. The protein contains at least one NAP domain which  
 CC has selective inhibitory activity for factor VIIa/TF. The specification  
 CC describes a method for screening an isolated protein at least one domain  
 CC for factor VIIa/TF selective inhibitory activity. The method comprises  
 CC determining the time to clotting effected by a concentration of the  
 CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo  
 CC activated partial thromboplastin time (APTT) assay; calculating  
 CC prolongation of clotting effected by the isolated protein in each of the  
 CC PT and APTT assay, with respect to a baseline clotting value for each  
 CC assay, where prolongation of clotting is calculated as fold elevation of  
 CC clotting time relative to a baseline clotting value, where a doubling of  
 CC clotting time is deemed a two-fold elevation; and calculating a PT to  
 CC APTT prolongation ratio, where a ratio at least one is indicative of  
 CC factor VIIa/TF inhibitory activity. The method is useful for determining  
 CC if a protein has factor VIIa/TF inhibitory activity  
 XX Sequence 84 AA;  
 Query Match 93.8%; Score 30; DB 2; Length 84;  
 Best Local Similarity 100.0%; Pred. No. 81;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 GFYRN 7  
 Db 56 GFYRN 60

Search completed: September 24, 2004, 07:34:16  
 Job time : 17.9474 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 07:28:33 ; Search time 4.27368 Seconds  
(without alignments)  
157.555 Million cell updates/sec

Title: US-09-498-556C-79

Perfect score: 32

Sequence: 1 XXGFYRN 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283356

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:\*\*

1: PIR1:\*\*

2: PIR2:\*\*

3: PIR3:\*\*

4: PIR4:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	93.8	77	2 B85684	unknown protein en
2	30	93.8	134	2 S76180	hypothetical prote
3	30	93.8	172	2 B87703	conserved hypotet
4	30	93.8	175	2 S73496	hypoxanthine-guan
5	30	93.8	175	2 F64250	hypoxanthine-guan
6	30	93.8	223	2 AC2302	endonuclease III
7	30	93.8	226	2 C70790	probable Endonucle
8	30	93.8	237	2 G84299	hypothetical prote
9	30	93.8	361	2 AG0169	phosphoserine tran
10	30	93.8	362	2 C82572	phosphoserine amin
11	30	93.8	408	2 T25524	hypothetical prote
12	30	93.8	415	2 T45716	hypothetical prote
13	30	93.8	465	2 A96553	probable myrosinase
14	30	93.8	481	2 S50053	tryptophan-tRNA li
15	30	93.8	533	1 YRWSCS	monophenol monooxy
16	30	93.8	747	2 T16274	hypothetical prote
17	30	93.8	824	2 S40937	hypothetical prote
18	30	93.8	900	2 S70630	xeroderma pigmento
19	30	93.8	940	2 S44345	AcRB/ACR/D/ACRF fam
20	30	93.8	1041	2 C87645	probable transport
21	30	93.8	1564	2 S55517	hypothetical prote
22	27	84.4	36	2 H70251	hypothetical prote
23	27	84.4	60	1 F70187	ribosomal protein
24	27	84.4	199	2 T39498	hypothetical prote
25	27	84.4	201	2 A86636	hypothetical prote
26	27	84.4	234	2 A84993	purine-nucleoside
27	27	84.4	258	2 T22092	hypothetical prote
28	27	84.4	275	2 A69413	conserved hypotet
29	27	84.4	299	2 F64491	hypothetical prote

30 27 84.4 318 2 B36972  
31 27 84.4 318 2 S33433  
32 27 84.4 340 2 D69414  
33 27 84.4 340 2 C75004  
34 27 84.4 353 2 E64376  
35 27 84.4 353 2 F69193  
36 27 84.4 426 2 F96994  
37 27 84.4 461 2 E95152  
38 27 84.4 463 2 H85745  
39 27 84.4 463 2 C90873  
40 27 84.4 465 2 D64881  
41 27 84.4 465 2 D90866  
42 27 84.4 465 2 B85752  
43 27 84.4 465 2 AC0659  
44 27 84.4 471 1 A41706  
45 27 84.4 471 1 S62590

## ALIGNMENTS

### RESULT 1

B85684  
unknown protein encoded by prophage CP-933C [imported] - Escherichia coli (strain O157:H7)  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C:Accession: B85684  
R:Perina, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: B85684  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-77 <STO>  
A:Cross-references: GB:AE005174; NID:g12514756; PIDN:AA055934.1; GSPDB:GNO0145; UMGP:Z1  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: Z1836

Query Match 93.8%; Score 30; DB 2; Length 77;  
Best Local Similarity 100.0%; Pred. No. 9.8;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
DB 15 GFYRN 19

### RESULT 2

S76180  
hypothetical protein - Synecocystis sp. (strain PCC 6803)  
C:Species: Synecocystis sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
C:Accession: S76180  
R:Kareko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;  
c, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
DNA Res. 3, 103-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis  
s.  
A:Reference number: S74322; MUID:97061201; PMID:8905231  
A:Accession: S76180  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-134 <KAN>  
A:Cross-references: EMBL:D80914; GB:AB001339; NID:g1653477; PIDN:BA018439.1; PID:d101917  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 93.8%; Score 30; DB 2; Length 134;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
|||||  
Db 27 GFYRN 31

## RESULT: 3

B87703  
C:Species: Caulobacter crescentus  
C:Title: conserved hypothetical protein CC3660 [imported] - Caulobacter crescentus  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C:Accession: B87703  
R:Niernman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of *Caulobacter crescentus*.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: B87703  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-172 <STO>  
A:Cross-references: GB:AE005673; NID:gl3425418; PIDN:AAK25622.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC3660

Query Match 93.8%; Score 30; DB 2; Length 172;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
|||||  
Db 73 GFYRN 77

## RESULT: 4

S73496  
hypoxanthine-guanine phosphoribosyltransferase hpt - Mycoplasma pneumoniae (strain ATCC  
N:Alternate names: Hypothetical protein K05\_orf175  
C:Species: Mycoplasma pneumoniae  
A:Variety: ATCC 29342  
C:Date: 27-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 07-Dec-1999  
C:Accession: S73496  
R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.  
Nucleic Acids Res. 24, 4420-4449, 1996  
A:Title: Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*  
A:Reference number: S73327; MUID:97105885; PMID:8948633  
A:Accession: S73496  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-175 <HIM>  
A:Cross-references: EMBL:AE000018; GB:U00089; NID:gl673827; PIDN:AB95818.1; PID:gl67383  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996  
C:Genetics:  
A:Gene: hpt  
A:Genetic code: SGC3  
C:Superfamily: hypoxanthine phosphoribosyltransferase

Query Match 93.8%; Score 30; DB 2; Length 175;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
|||||  
Db 159 GFYRN 163

## RESULT: 5

F64250  
hypoxanthine-guanine phosphoribosyltransferase - Mycoplasma genitalium  
C:Species: Mycoplasma genitalium  
C:Date: 17-Nov-1995 #sequence\_revision 17-Nov-1995 #text\_change 07-Dec-1999  
C:Accession: F64250

R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;  
M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.N.  
C.A.; Venter, J.C.  
Science 270, 397-403, 1995  
A:Title: The minimal gene complement of *Mycoplasma genitalium*.  
A:Reference number: A64200; MUID:96026346; PMID:7569993  
A:Accession: F64250  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-175 <TIGR>  
A:Cross-references: GB:U39733; GB:L43967; NID:gl046177; PIDN:AA01648.1; PID:gl046178; TJ  
A:Experimental source: strain G-37  
A:Genetic code: SGC3  
C:Superfamily: hypoxanthine phosphoribosyltransferase

Query Match 93.8%; Score 30; DB 2; Length 175;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
|||||  
Db 159 GFYRN 163

## RESULT: 6

AC2302  
endonuclease III [imported] - Nostoc sp. (strain PCC 7120)  
C:Species: Nostoc sp. PCC 7120  
A:Note: Nostoc sp. Strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C:Accession: AC2302  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,  
N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing *Cyanobacterium* *Anabaena*  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AC2302  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-223 <KUR>  
A:Cross-references: GB:BA000019; PIDN:BA075669.1; PID:gl7133104; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: nth  
C:Superfamily: apurinic/apyrimidinic endonuclease III

Query Match 93.8%; Score 30; DB 2; Length 223;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
|||||  
Db 89 GFYRN 93

## RESULT: 7

C70790  
probable Endonuclease III - *Mycobacterium tuberculosis* (strain H37RV)  
C:Species: *Mycobacterium tuberculosis*  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C:Accession: C70790  
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.  
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.  
Rajandream, M.A.; Rogers, J.; Ruter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: C70790  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-226 <COL>

A;Cross-references: GB:AL022121; GB:AL123456; NID:g3261559; PIDN:CAA17996.1; PID:g296009  
A;Experimental source: strain H37RV  
C;Genetics:  
A;Gene: nth  
C;Superfamily: apurinic/aprimidinic endonuclease III  
C;Keywords: 4Fe-4S; iron-sulfur protein; metalloprotein  
F:179,186,189,195/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 93.8%; Score 30; DB 2; Length 226;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7  
Db 72 GFYRN 76

RESULT 8  
G84299  
Hypothetical protein Vng1457c [imported] - Halobacterium sp. NRC-1  
C;Species: Halobacterium sp. NRC-1  
C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C;Accession: G84299  
R;NG, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Laskey, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabil  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li  
A;Title: Genome sequence of Halobacterium species NRC-1.  
A;Reference number: A84160; MUID:20504483; PMID:11016950  
A;Accession: G84299  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-237 <STO>  
A;Cross-references: GB:AE004437; NID:g10580953; PIDN:AAG19763.1; GSPDB:GN00138  
C;Genetics:  
A;Gene: Vng1457C

Query Match 93.8%; Score 30; DB 2; Length 237;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7  
Db 191 GFYRN 195

RESULT 9  
AG0169  
Phosphoserine transaminase (EC 2.6.1.52) [imported] - Yersinia pestis (strain CO92)  
C;Species: Yersinia pestis  
C;Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 27-Nov-2001  
C;Accession: AG0169  
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,  
Nature 413, 523-527, 2001  
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A;Reference number: AB0001; MUID:21470413; PMID:11586360  
A;Accession: AG0169  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-361 <KUR>  
A;Cross-references: GB:AL590842; PIDN:CAC90218.1; PID:g15979439; GSPDB:GN00175  
C;Genetics:  
A;Gene: serC  
C;Superfamily: phosphoserine aminotransferase  
C;Keywords: aminotransferase

Query Match 93.8%; Score 30; DB 2; Length 361;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7  
Db 283 GFYRN 287

RESULT 10  
C82572  
Phosphoserine aminotransferase XF2326 [imported] - Xylella fastidiosa (strain 9a5c)  
C;Species: Xylella fastidiosa  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Sep-2000  
C;Accession: C82572  
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen  
Nature 406, 151-157, 2000  
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A;Reference number: A82515; MUID:20365717; PMID:10910347  
A;Note: for a complete list of authors see reference number A59328 below  
A;Accession: C82572  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-362 <SIM>  
A;Cross-references: GB:AE004043; GB:AE003849; NID:g9107486; PIDN:AAF85125.1; GSPDB:GN001  
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A  
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H  
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000  
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Praga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E  
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.;  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A;Reference number: A59328  
A;Cross-references: GB:AE004043; GB:AE003849; NID:g9107486; PIDN:AAF85125.1; GSPDB:GN001  
C;Genetics:  
A;Gene: XF2326  
C;Superfamily: phosphoserine aminotransferase

Query Match 93.8%; Score 30; DB 2; Length 362;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7  
Db 284 GFYRN 288

RESULT 11  
T25524  
Hypothetical protein C06A5.9 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C;Accession: T25524  
R;Davidson, S.; Wohlmann, P.  
submitted to the EMBL Data Library, April 1997  
A;Description: The sequence of C. elegans cosmid C06A5.  
A;Reference number: Z20044  
A;Accession: T25524  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-408 <DAV>  
A;Cross-references: EMBL:U97193; PIDN:AAB52442.1; GSPDB:GN00019; CESP:C06A5.9  
A;Experimental source: strain Bristol N2; clone C06A5  
C;Genetics:  
A;Gene: CESP:C06A5.9  
A;Map position: 1  
A;Introns: 3/1; 27/3; 86/1; 154/1; 201/3; 240/3; 268/3; 297/3

Query Match 93.8%; Score 30; DB 2; Length 408;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
DB 229 GFYRN 233

RESULT 12  
T46716  
hypothetical protein L4326.05 [imported] - Leishmania major  
C:Species: Leishmania major  
C>Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 18-Feb-2000  
C/Accession: T46716  
R/Volckaert, G.; Ivens, A.C.; Lawson, D.; Quail, M.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, December 1999  
A:Reference number: 223137  
A/Accession: T46716  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-415 <VOL>  
A/Cross-references: EMBL:AL121861; PIDN:CAB58381.1  
A:Experimental source: strain Friedlin  
C/Genetics:  
A>Note: L4326.05

Query Match 93.8%; Score 30; DB 2; Length 415;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
DB 325 GFYRN 329

RESULT 13  
A96553  
probable myrosinase precursor 5323-50499 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001  
C/Accession: A96553  
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;  
ansen, N.P.; Hughes, B.; Huizar, L.  
Nature 406, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A96141; MUID:21016719; PMID:11130712  
A/Accession: A96553  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-465 <STO>  
A/Cross-references: GB:AE005173; NID:g10092358; PIDN:AAG12767.1; GSPDB:GN00141  
C/Genetics:  
A:Gene: P5D21.17  
A:Map position: 1  
C:Superfamily: Agrobacterium beta-glucosidase

Query Match 93.8%; Score 30; DB 2; Length 465;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
DB 331 GFYRN 335

RESULT 14  
S50053  
tryptophan-tRNA ligase (EC 6.1.1.2) alpha-2 chain - mouse  
C:Species: Mus musculus (house mouse)

C>Date: 07-May-1995 #sequence\_revision 10-Nov-1995 #text\_change 03-Jun-2002  
C/Accession: S50053; S50052; I49391; S31461; S31462  
R/Pajot, B.; Sarger, C.; Bonnet, J.; Garret, M.  
J. Mol. Biol. 242, 599-603, 1994  
A:Title: An alternative splicing modifies the C-terminal end of tryptophanyl-tRNA synthetase  
A:Reference number: S50052; MUID:95018226; PMID:7932716  
A/Accession: S50053  
A:Molecule type: mRNA  
A:Residues: 1-481 <PAJ>  
A/Cross-references: EMBL:X69657; NID:g55437; PIDN:CAA49348.1; PID:g55438  
A:Genetics: LSF  
A>Note: Intron position was determined by sequencing of genomic DNA  
A/Accession: S50052  
A:Molecule type: mRNA  
A:Residues: 1-475 <PAW>  
A/Cross-references: EMBL:X69656; NID:g55435; PIDN:CAA49347.1; PID:g55436  
R/Kisselev, L.L.  
Biochimie 75, 1027-1039, 1993  
A:Title: Mammalian tryptophanyl-tRNA synthetases.  
A:Reference number: I49391; MUID:94257729; PMID:7515282  
A/Accession: I49391  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-481 <RES>  
A/Cross-references: EMBL:X69657; NID:g55437; PIDN:CAA49348.1; PID:g55438  
C/Genetics: <LSF>  
A:Gene: WRS  
A:Introns: 475/2  
A>Note: the list of introns may be incomplete; clone W13  
C/Genetics: <SSF>  
A:Gene: WRS  
A>Note: clone S5  
C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology  
C/Keywords: alternative splicing; aminoacyl-tRNA synthetase; ligase; protein biosynthesis  
F1-481/Product: tryptophan-tRNA ligase alpha-2 chain long splice form #status predicted  
F1-475/Product: tryptophan-tRNA ligase alpha-2 chain short splice form #status predicted  
F1-23-68/Domain: amino acid-tRNA ligase repeat homology <ATL>

Query Match 93.8%; Score 30; DB 2; Length 481;  
Best Local Similarity 100.0%; Pred. No. 62;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
DB 250 GFYRN 254

RESULT 15  
YRMSGS  
monophenol monooxygenase (EC 1.14.18.1) precursor [validated] - mouse  
A/Alternate names: cresolase; monophenol oxidase; phenolase; tyrosinase  
C/Species: Mus musculus (house mouse)  
C>Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 20-Apr-2000  
C/Accession: A27711; A60778; B32429; B32429; S01170; S02278; S15753; I49736  
R/Kwon, B.S.; Wakulchik, M.; Haq, A.K.; Halaban, R.; Kestler, D.  
Biochem. Biophys. Res. Commun. 153, 1301-1309, 1988  
A:Title: Sequence analysis of mouse tyrosinase cDNA and the effect of melanotropin on its  
A:Reference number: A27711; MUID:88288910; PMID:3134020  
A/Accession: A27711  
A:Molecule type: mRNA  
A:Residues: 1-533 <KWO>  
A/Cross-references: GB:M20234; NID:g202247; PIDN:AAA40516.1; PID:g202248  
A:Experimental source: Cloudman S-91 melanoma cells  
R/Kwon, B.S.; Haq, A.K.; Wakulchik, M.; Kestler, D.; Barton, D.E.; Francke, U.; Lamoreux,  
J. Invest. Dermatol. 93, 589-594, 1989  
A:Title: Isolation, chromosomal mapping, and expression of the mouse tyrosinase gene.  
A:Reference number: A60778; MUID:90010220; PMID:2507645  
A/Accession: A60778  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-273 <KW2>  
A:Experimental source: BALB/c

R;Terao, M.; Tabe, L.; Garattini, E.; Sartori, D.; Studer, M.; Mintz, B.  
Biochem. Biophys. Res. Commun. 159, 848-853, 1989  
A;Title: Isolation and characterization of variant cDNAs encoding mouse tyrosinase.  
A;Reference number: A32429; MUID:89193679; PMID:2494997  
A;Accession: A32429  
A;Molecule type: mRNA  
A;Residues: 1-102, 'C', 104-345, 'G', 347-533 <TER>  
A;Cross-references: GB:M24560; NID:G202249; PIDN:AAA40517.1; PID:G202250  
A;Accession: B32429  
A;Molecule type: mRNA  
A;Residues: 1-77;155-345, 'G', 347-533 <TE2>  
A;Cross-references: GB:M24560  
A;Experimental source: B16 melanoma cells  
R;Mueller, G.; Ruppert, S.; Schmid, E.; Schuetz, G.  
EMBO J. 7, 2723-2730, 1988  
A;Title: Functional analysis of alternatively spliced tyrosinase gene transcripts.  
A;Reference number: S01170; MUID:89030636; PMID:3141148  
A;Accession: S01170  
A;Molecule type: mRNA  
A;Residues: 1-102, 'C', 104-263, 'I', 265-345, 'G', 347-533 <MUE>  
A;Cross-references: GB:X12782; NID:G55061; PIDN:CAA3273.1; PID:G55062  
R;Yamamoto, H.; Takeuchi, S.; Kudo, T.; Makino, K.; Nakata, A.; Shinoda, T.; Takeuchi, T.  
Jpn. J. Genet. 62, 271-274, 1987  
A;Title: Cloning and sequencing of mouse tyrosinase cDNA.  
A;Reference number: S02278  
A;Accession: S02278  
A;Molecule type: mRNA  
A;Residues: 1-102, 'C', 104-263, 'I', 265-345, 'G', 347-448 <YAM>  
A;Cross-references: EMBL:X12782  
A;Note: part of this sequence was confirmed by protein sequencing  
R;Shibahara, S.; Okinaga, S.; Tomita, Y.; Takeda, A.; Yamamoto, H.; Sato, M.; Takeuchi, H.  
Eur. J. Biochem. 189, 455-461, 1990  
A;Title: A point mutation in the tyrosinase gene of BALB/c albino mouse causing the cyst  
A;Reference number: S15753; MUID:90249393; PMID:2110899  
A;Accession: S15753  
A;Status: translation not shown  
A;Molecule type: DNA  
A;Residues: 1-13 <SHI>  
A;Cross-references: EMBL:X51743; NID:G55057; PIDN:CAA36033.1; PID:G55058  
R;Kwon, B.S.; Halaban, R.; Chintamaneni, C.  
Biochem. Biophys. Res. Commun. 161, 252-260, 1989  
A;Title: Molecular basis of mouse Himalayan mutation.  
A;Reference number: I49736; MUID:89273644; PMID:2567165  
A;Accession: I49736  
A;Status: preliminary; translated from GE/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-39, 'I', 41-102, 'C', 104-196, 'Q', 198-345, 'G', 347-419, 'R', 421-533 <RES>  
A;Cross-references: GB:M26729; NID:G193845; PIDN:AAA37806.1; PID:G309296  
C;Comment: This cell-specific oxidase is a glycoprotein containing two Cu per enzyme; it  
reactions in the formation of pigments such as melanins and other polypheolic compound  
C;Genetics:  
A;Gene: Tyr1  
A;Map position: 7  
C;Superfamily: monophenol monooxygenase  
C;Keywords: albinism; alternative splicing; copper; glycoprotein; melanin biosynthesis;  
F;1-18/domain: signal sequence #status predicted <SIG>  
F;19-533/Product: monophenol monooxygenase #status predicted <MAT>  
F;474-497/Domain: transmembrane #status predicted <TM>  
F;86,111,161,230,337/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 93.8%; Score 30; DB 1; Length 533;  
Best Local Similarity 100.0%; Pred. No. 69;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 GFYRN 7  
|||||  
DB 458 GFYRN 462





DR TIGR; MG458; --  
DR InterPro; IPR005904; Hxn\_phospho trans.  
DR InterPro; IPR002375; Pr/PY\_rd\_transf.  
DR InterPro; IPR000836; PRTransf.ase.  
DR Pfam; PF00156; Priboyltran; 1.  
DR TIGRFAMs; TIGR01203; HGPrtase; 1.  
DR PROSITE; PS00103; PUR\_PYR\_TRANSF.ER; 1.  
DR Transf.ase; Glycosyltransferase; Purine salvage; Metal-binding;  
KW Magnesium; Complete proteome.  
FT METAL 156 MAGNESIUM 1 (BY SIMILARITY).  
SQ SEQUENCE 175 AA; 19759 MW; 9647332FF3E18E0 CRC64;  
  
Query Match 93.8%; Score 30; DB 1; Length 175;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 3 GFYRN 7  
DB 159 GFYRN 163  
  
RESULT 2  
HPRT MYCPN STANDARD; PRT; 175 AA.  
ID HPRT MYCPN STANDARD; PRT; 175 AA.  
AC P75119;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hypoxanthine-guanine phosphoribosyltransferase (EC 2.4.2.8) (HGPR)  
DE (HGPRase).  
GN HPT OR MPN672 OR MP170.  
OS Mycoplasma pneumoniae.  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=21104;  
RN [1]  
SEQUENCE FROM N.A.  
RC STRAIN=ATCC 29342 / M129;  
RX MEDLINE=97105885; PubMed=8948633;  
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,  
RA Herrmann R.;  
RT "complete sequence analysis of the genome of the bacterium Mycoplasma  
RT pneumoniae.";  
RL Nucleic Acids Res. 24:4420-4445 (1996).  
CC -!- CATALYTIC ACTIVITY: IMP + diphosphate = hypoxanthine + 5-phospho-  
CC -!- CATALYTIC ACTIVITY: GMP + diphosphate = guanine + 5-phospho-  
CC -!- alpha-D-ribose 1-diphosphate.  
CC -!- alpha-D-ribose 2-diphosphate.  
CC -!- COFACTOR: Binds 2 magnesium ions per subunit. One of the ions does  
CC not make direct protein contacts (By similarity).  
CC -!- PATHWAY: Purine salvage.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- SIMILARITY: Belongs to the purine/pyrimidine  
CC phosphoribosyltransferase family.  
  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
  
CC EMBL; AB000018; AAB95818.1; --  
CC F1R; S73496; S73496.  
CC HSP; P00492; 1HNP.  
DR InterPro; IPR005904; Hxn\_phospho trans.  
DR InterPro; IPR002375; Pr/PY\_rd\_transf.  
DR InterPro; IPR000836; PRTransf.ase.  
DR Pfam; PF00156; Priboyltran; 1.  
DR TIGRFAMs; TIGR01203; HGPrtase; 1.  
DR PROSITE; PS00103; PUR\_PYR\_TRANSF.ER; 1.  
DR Transf.ase; Glycosyltransferase; Purine salvage; Metal-binding;  
KW Magnesium; Complete proteome.  
KW

FT METAL 156 MAGNESIUM 1 (BY SIMILARITY).  
SQ SEQUENCE 175 AA; 19620 MW; 1E5A9FRA01D69854 CRC64;  
  
Query Match 93.8%; Score 30; DB 1; Length 175;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 3 GFYRN 7  
DB 159 GFYRN 163  
  
RESULT 3  
END3 MYCTU STANDARD; PRT; 245 AA.  
ID END3 MYCTU STANDARD; PRT; 245 AA.  
AC O69642;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Endonuclease III (EC 4.2.99.18) (DNA-(apurinic or apyrimidinic site)  
DE lyase).  
GN NTH OR RV3674C OR MT3775 OR MTV025.022C OR MB3698C.  
OS Mycobacterium tuberculosis, and  
OS Mycobacterium bovis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773, 1765;  
RN [1]  
SEQUENCE FROM N.A.  
RC SPECIES=M.tuberculosis; STRAIN=H37Rv;  
RX MEDLINE=98295987; PubMed=9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,  
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
RT "Deciphering the biology of Mycobacterium tuberculosis from the  
RT complete genome sequence.";  
RL Nature 393:537-544 (1998).  
RN [2]  
SEQUENCE FROM N.A.  
RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;  
RX MEDLINE=2206494; PubMed=12218036;  
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,  
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,  
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;  
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and  
RT laboratory strains.";  
RL J. Bacteriol. 184:5479-5490 (2002).  
RN [3]  
SEQUENCE FROM N.A.  
RC SPECIES=M.bovis; STRAIN=AF2122/97;  
RX MEDLINE=22709107; PubMed=12788972;  
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,  
RA Pryor M., Duthey S., Grondin S., Lacroix C., Monsemp C., Simon S.,  
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,  
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;  
RT "The complete genome sequence of Mycobacterium bovis.";  
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).  
CC -!- FUNCTION: Has both an apurinic and/or apyrimidinic endonuclease  
CC activity and a DNA N-glycosylase activity. Incises damaged DNA at  
CC cytosines, thymines and guanines. Acts on a damaged strand, 5'  
CC from the damaged site (By similarity).  
CC -!- CATALYTIC ACTIVITY: The C-O-P bond 3' to the apurinic or  
CC apyrimidinic site in DNA is broken by a beta-elimination reaction,  
CC leaving a 3'-terminal unsaturated sugar and a product with a  
CC terminal 5'-phosphate.  
CC -!- COFACTOR: Binds a 4Fe-4S cluster which is not important for the

CC catalytic activity, but which is probably involved in the proper  
CC positioning of the enzyme along the DNA strand (By similarity).  
CC -1- SIMILARITY: Belongs to the nth/mutY family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AL022121; CAJ17996.1; ALT\_INIT.  
CC EMBL; AE007175; AAK48142.1; ALT\_INIT.  
CC EMBL; BX248346; CAD95884.1; -  
CC HSSP; P20625; 2ABK.  
CC TIGR; MT3775; -  
CC TubercuList; RV3674C; -  
CC InterPro; IPR003265; Endo\_3c.  
CC InterPro; IPR004035; EndoIII\_FCL.  
CC InterPro; IPR004036; EndoIII\_HhH.  
CC InterPro; IPR003651; Fes\_bind.  
CC InterPro; IPR004445; HhH.  
CC InterPro; IPR003583; HhH\_1.  
CC InterPro; IPR005759; Nth.  
CC Pfam; PF00730; HhH-GPD; 1.  
CC Pfam; PF06633; HhH; 1.  
CC SMART; SM00478; ENDO3c; 1.  
CC SMART; SM00525; FES; 1.  
CC SMART; SM00278; HhH1; 1.  
CC TIGRFAMs; TIGR01083; nth; 1.  
CC PROSITE; PS00764; ENDONUCLEASE\_III\_1; 1.  
CC PROSITE; PS01155; ENDONUCLEASE\_III\_2; 1.  
CC Hydrolase; Nuclease; Iron-sulfur; 4Fe-4S; Complete proteome.  
KW Glycosidase; Lyase; Iron-sulfur; 4Fe-4S; Multifunctional enzyme; DNA repair;  
FT METAL 198 198 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
FT METAL 205 205 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
FT METAL 208 208 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
FT METAL 214 214 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
SQ SEQUENCE 245 AA; 27030 MW; 2B6D16195DD090DE CRC64;  
Query Match 93.8%; Score 30; DB 1; Length 245;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 GFYRN 7  
Db 91 GFYRN 95  
RESULT 4  
ID UVEN MICLU STANDARD; PRT; 279 AA.  
AC P46303;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Ultraviolet N-glycosylase/AP lyase (UV-endonuclease) (Pyrimidine dimer  
DN glycosylase).  
GN PDG.  
OS Micrococcus luteus (Micrococcus lysodeikticus).  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Micrococineae; Micrococcaceae; Micrococcus.  
OX NCBI\_TaxID=1270;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-35.  
RC STRAIN=ATCC 4698;  
RX MEDLINE=96007490; PubMed=7559510;  
RA Piersen C.E., Prince M.A., Augustine M.L., Dodson M.L., Lloyd R.S.;  
RT "Purification and cloning of Micrococcus luteus ultraviolet  
RT endonuclease, an N-glycosylase/abasic lyase that proceeds via an  
RT imino enzyme-DNA intermediate."  
RL J. Biol. Chem. 270:23475-23484(1995).

CC -1- FUNCTION: Has both, an apurinic and/or apyrimidinic endonuclease  
CC activity and a DNA N-glycosylase activity. Initiates repair at  
CC cis-syn pyrimidine dimers. Proceeds via an imino enzyme-DNA  
CC intermediate.  
CC -1- MISCELLANEOUS: Readthrough of the terminator UAG occurs between  
CC codons for Gly-268 and Ala-270. Two forms of 31 kDa and 32 kDa  
CC have been detected.  
CC -1- SIMILARITY: Belongs to the nth/mutY family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U22181; AAA86508.1; ALT\_TERM.  
CC HSSP; P20625; 2ABK.  
CC InterPro; IPR003265; Endo\_3c.  
CC InterPro; IPR004035; EndoIII\_FCL.  
CC InterPro; IPR004036; EndoIII\_HhH.  
CC InterPro; IPR003651; Fes\_bind.  
CC InterPro; IPR004445; HhH.  
CC InterPro; IPR003583; HhH\_1.  
CC InterPro; IPR005759; Nth.  
CC Pfam; PF00730; HhH-GPD; 1.  
CC Pfam; PF06633; HhH; 1.  
CC SMART; SM00478; ENDO3c; 1.  
CC SMART; SM00525; FES; 1.  
CC SMART; SM00278; HhH1; 1.  
CC TIGRFAMs; TIGR01083; nth; 1.  
CC PROSITE; PS00764; ENDONUCLEASE\_III\_1; 1.  
CC PROSITE; PS01155; ENDONUCLEASE\_III\_2; 1.  
CC Hydrolase; Nuclease; Endonuclease; DNA repair; Glycosidase;  
KW Iron-sulfur; 4Fe-4S.  
FT CHAIN 1 268 UV ENDONUCLEASE 31 kDa FORM.  
FT CHAIN 1 279 UV ENDONUCLEASE 32 kDa FORM (PROBABLE).  
FT METAL 203 203 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
FT METAL 210 210 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
FT METAL 213 213 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
FT METAL 219 219 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
SQ SEQUENCE 279 AA; 30469 MW; C385369A1827C005 CRC64;  
Query Match 93.8%; Score 30; DB 1; Length 279;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 GFYRN 7  
Db 95 GFYRN 99  
RESULT 5  
ID SERC\_XANAC STANDARD; PRT; 361 AA.  
AC Q8PL77;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Phosphoserine aminotransferase (EC 2.6.1.52) (PSAT).  
GN SERC OR XAC1648.  
OS Xanthomonas axonopodis (pv. citri).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
OC Xanthomonadaceae; Xanthomonas.  
OX NCBI\_TaxID=92829;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=306 / ATCC 13902 / XV 101;  
RX MEDLINE=2202145; PubMed=12024217;  
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,  
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.P.,  
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,

RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,  
RA Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,  
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,  
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,  
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,  
RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,  
RA Martins E.C., Meidanis J., Menck C.F.M., Oliveira M.C., Oliveira V.R.,  
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,  
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,  
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,  
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,  
RA Setubal J.C., Kitajima J.P.;  
RT "Comparison of the genomes of two Xanthomonas pathogens with differing  
RT host specificities.";  
RL Nature 417:459-463 (2002).  
CC -!- CATALYTIC ACTIVITY: O-phospho-L-serine + 2-oxoglutarate = 3-  
CC phosphonocopyruvate + L-glutamate.  
CC -!- COFACTOR: Pyridoxal phosphate (By similarity).  
CC -!- PATHWAY: Required both in major phosphorylated pathway of serine  
CC biosynthesis and in the biosynthesis of pyridoxine (By  
CC similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: Belongs to class-V of pyridoxal-phosphate-dependent  
CC aminotransferases.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AE011797; AAM36516.1; -;  
CC HAMAP; MF\_00160; -; 1.  
CC InterPro; IPR000192; Aminotransf. V.  
CC InterPro; IPR003248; Pser\_aminotransf.  
CC Pfam; PF00266; aminotran\_5; 1.  
CC ProDom; PD001544; Pser\_aminotransf; 1.  
CC TIGRFAMS; TIGR01364; serC1; 1.  
CC DR PROSITE; PS00595; AA\_TRANSF\_CLASS\_5; 1.  
CC Serine biosynthesis; Pyridoxine biosynthesis; Transferase;  
CC Aminotransferase; Pyridoxal phosphate; Complete proteome.  
CC BINDING 196 196 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
CC FT SEQUENCE 361 AA; 38680 MW; 75EC2B319C493982 CRC64;  
CC -----  
CC Query Match 93.8%; Score 30; DB 1; Length 361;  
CC Best Local Similarity 100.0%; Pred. No. 28;  
CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CC -----  
CC QY 3 GFYRN 7  
CC Db 283 GFYRN 287  
CC -----  
CC RESULT 6  
CC SERC\_XANCP  
CC ID SERC\_XANCP STANDARD; PRT; 361 AA.  
CC AC Q8PA97;  
CC DT 10-OCT-2003 (Rel. 42, Created)  
CC DT 10-OCT-2003 (Rel. 42, Last sequence update)  
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)  
CC DE Phosphoserine aminotransferase (EC 2.6.1.52) (PSAT).  
CC GN SERC OR XCCL589.  
CC OS Xanthomonas campestris (pv. campestris).  
CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
CC OC Xanthomonadaceae; Xanthomonas.  
CC OX NCBI\_TaxID=340;  
CC OX [1]  
CC RN Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
CC RP SEQUENCE FROM N.A.  
CC RC STRAIN=ATCC 33913 / NCPPB 528;  
CC RX MEDLINE=2202145; PubMed=12024217;  
CC RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,

RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,  
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,  
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,  
RA Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,  
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,  
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,  
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,  
RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,  
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,  
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,  
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,  
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,  
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,  
RA Setubal J.C., Kitajima J.P.;  
RT "Comparison of the genomes of two Xanthomonas pathogens with differing  
RT host specificities.";  
RL Nature 417:459-463 (2002).  
CC -!- CATALYTIC ACTIVITY: O-phospho-L-serine + 2-oxoglutarate = 3-  
CC phosphonocopyruvate + L-glutamate.  
CC -!- COFACTOR: Pyridoxal phosphate (By similarity).  
CC -!- PATHWAY: Required both in major phosphorylated pathway of serine  
CC biosynthesis and in the biosynthesis of pyridoxine (By  
CC similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: Belongs to class-V of pyridoxal-phosphate-dependent  
CC aminotransferases.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AE012260; AAM40884.1; -;  
CC HAMAP; MF\_00160; -; 1.  
CC InterPro; IPR000192; Aminotrans V.  
CC InterPro; IPR003248; Pser\_aminotransf.  
CC Pfam; PF00266; aminotran\_5; 1.  
CC ProDom; PD001544; Pser\_aminotransf; 1.  
CC TIGRFAMS; TIGR01364; serC1; 1.  
CC DR PROSITE; PS00595; AA\_TRANSF\_CLASS\_5; 1.  
CC Serine biosynthesis; Pyridoxine biosynthesis; Transferase;  
CC Aminotransferase; Pyridoxal phosphate; Complete proteome.  
CC BINDING 196 196 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
CC FT SEQUENCE 361 AA; 38759 MW; F0025F9E9BF77E65 CRC64;  
CC -----  
CC Query Match 93.8%; Score 30; DB 1; Length 361;  
CC Best Local Similarity 100.0%; Pred. No. 28;  
CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CC -----  
CC QY 3 GFYRN 7  
CC Db 283 GFYRN 287  
CC -----  
CC RESULT 7  
CC SERC\_YERPE  
CC ID SERC\_YERPE STANDARD; PRT; 361 AA.  
CC AC Q8ZGE4;  
CC DT 10-OCT-2003 (Rel. 42, Created)  
CC DT 10-OCT-2003 (Rel. 42, Last sequence update)  
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)  
CC DE Phosphoserine aminotransferase (EC 2.6.1.52) (PSAT).  
CC GN SERC OR YP01389 OR Y2784.  
CC OS Yersinia pestis.  
CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
CC OC Enterobacteriaceae; Yersinia.  
CC OX NCBI\_TaxID=632;  
CC OX [1]  
CC RN Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
CC RP SEQUENCE FROM N.A.  
CC RC STRAIN=CO-92 / Biovar Orientalis;  
CC RA

```

RX MEDLINE=21470413; PubMed=11596360;
RA ParKhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moulé S., Oyston P.C.F., Quail M.A., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
RA "Genome sequence of Yersinia pestis, the causative agent of plague.";
RA Nature 413:523-527(2001).
RN [12]
RP SEQUENCE FROM N.A.
RC STRAIN=KIMS / Biovar Mediaevalis;
RX MEDLINE=221337963; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.B., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Millies M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
EL J. Bacteriol. 184:4601-4611(2002).
CC -!- CATALYTIC ACTIVITY: O-phospho-L-serine + 2-oxoglutarate = 3-
CC phosphonoxypropylate + L-glutamate.
CC -!- COFACTOR: Pyridoxal phosphate (By similarity).
CC -!- PATHWAY: Required both in major phosphorylated pathway of serine
CC biosynthesis and in the biosynthesis of pyridoxine (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to class-v of pyridoxal-phosphate-dependent
CC aminotransferases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ414148; CAC90218.1; -.
DR EMBL; AE013881; AM86336.1; -.
DR PIR; AG0169; AG0169.
DR HAMAP; MF_00160; -. 1.
DR InterPro; IPR000192; Aminotrans V.
DR InterPro; IPR003248; Pser aminotransf.
DR Pfam; PF00266; aminotran 5; 1.
DR ProDom; PD001544; Pser aminotransf; 1.
DR TIGRFAMs; TIGR01364; serC 1; 1.
DR PROSITE; PS00595; AA_TRANSFER_CLASS_5; 1.
DR Serine biosynthesis; Pyridoxine biosynthesis; Transferase;
KW Serine biosynthesis; Pyridoxine biosynthesis; Complete proteome.
KW Aminotransferase; Pyridoxal phosphate; Complete proteome.
FT BINDING 197 197 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 361 AA; 40083 MW; 8827EF1419782D88 CRC64;

Query Match 93.8%; Score 30; DB 1; Length 361;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
Db 283 GFYRN 287
|||||
|||||

RESULT 8
SERC XYLF6 STANDARD; PRT; 362 AA.
AC Q9PBI9;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Phosphoserine aminotransferase (EC 2.6.1.52) (PSAT).
GN SERC OR XF2326.
OS Xylella fastidiosa.

```

```

OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9a5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.B.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Pacinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Frega J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Kiegr J.E., Kuranae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.B.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Silveira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshukano M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000)
CC -!- CATALYTIC ACTIVITY: O-phospho-L-serine + 2-oxoglutarate = 3-
CC phosphonoxypropylate + L-glutamate.
CC -!- COFACTOR: Pyridoxal phosphate (By similarity).
CC -!- PATHWAY: Required both in major phosphorylated pathway of serine
CC biosynthesis and in the biosynthesis of pyridoxine (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to class-v of pyridoxal-phosphate-dependent
CC aminotransferases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE004043; AAF85125.1; -.
DR PIR; C82572; C82572.
DR HSPF; P23721; 15JN.
DR HAMAP; MF_00160; -. 1.
DR InterPro; IPR000192; Aminotrans V.
DR InterPro; IPR003248; Pser aminotransf.
DR Pfam; PF00266; aminotran 5; 1.
DR ProDom; PD001544; Pser aminotransf; 1.
DR TIGRFAMs; TIGR01364; serC 1; 1.
DR PROSITE; PS00595; AA_TRANSFER_CLASS_5; FALSE_NEG.
DR Serine biosynthesis; Pyridoxine biosynthesis; Transferase;
KW Serine biosynthesis; Pyridoxine biosynthesis; Complete proteome.
KW Aminotransferase; Pyridoxal phosphate; Complete proteome.
FT BINDING 197 197 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 362 AA; 39615 MW; E0CE351A7A6276A4 CRC64;

Query Match 93.8%; Score 30; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7

```

```

Db      284 GFYRN 288
|||||
RESULT 9
SERC_XYLFT STANDARD; PRT; 362 AA.
AC Q87BU0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Phosphoserine aminotransferase (EC 2.6.1.52) (PSAT).
GN SERC OR FDI358;
OS Ralstonia solanacearum (strain Temeculal / ATCC 700964).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=183190;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22421331; PubMed=12533478;
RA Van Sluys M.A., de Oliveira M.C., Monteiro-Vitorello C.B.,
RA Miyaki C.Y., Furlan L.R., Camargo L.E.A., da Silva A.C.R., Moon D.H.,
RA Takita M.A., Lemos E.G.M., Machado M.A., Ferro M.I.T., da Silva F.R.,
RA Goldman M.H.S., Goldman G.H., Lemos M.V.F., El-Dorri H., Tsai S.M.,
RA Carver H., Carraro D.M., de Oliveira R.C., Nunes L.R., Siqueira W.J.,
RA Coutinho L.L., Kimura E.T., Ferro E.S., Harakava R., Kuranae E.E.,
RA Marino C.L., Gigliotti E., Abreu I.B., Alves L.M.C., do Amaral A.M.,
RA Baia G.S., Blanco S.R., Brito M.S., Cannavan F.S., Celestino A.V.,
RA da Cunha A.P., Fenille R.C., Ferro J.A., Formighieri E.F., Kishi L.T.,
RA Leoni S.G., Oliveira A.R., Rosa V.E. Jr., Sassaki F.T., Sena J.A.D.,
RA de Souza A.A., Truffi D., Tsukumo F., Yanai G.M., Zaros L.G.,
RA Civerolo E.L., Simpson A.J.G., Almeida N.F. Jr., Setubal J.C.,
RA Kitajima J.P.;
RT Comparative analyses of the complete genome sequences of Pierce's
RT disease and citrus variegated chlorosis strains of Xylella
RT fastidiososa."
RL J. Bacteriol. 185:1018-1026(2003).
CC -!- CATALYTIC ACTIVITY: O-phospho-L-serine + 2-oxoglutarate = 3-
CC phosphonooxypyruvate + L-glutamate.
CC -!- COFACTOR: Pyridoxal phosphate (By similarity).
CC -!- PATHWAY: Required both in major phosphorylated pathway of serine
CC biosynthesis and in the biosynthesis of pyridoxine (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to class-V of pyridoxal-phosphate-dependent
CC aminotransferases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF012558; AAC29205.1; -.
CC HAMAP; MF_00160; -.
CC InterPro; IPR000192; Aminotrans V.
CC InterPro; IPR003248; Pser_aminTransf.
CC Pfam; PF00266; aminotran_5; 1.
CC ProDom; PD001544; Pser_amintransf; 1.
CC TIGRFAMs; TIGR01364; serC.1; 1.
CC PROSITE; PS00595; AA_TRANSFER_CLASS_5; FALSE NEG.
CC Serine biosynthesis; Pyridoxine biosynthesis; Transferase;
CC Aminotransferase; Pyridoxal phosphate; Complete proteome.
CC BINDING 197 197 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC SEQUENCE 362 AA; 39502 MW; D40D59FB8B1D91C9 CRC64;
CC
CC Query Match 93.8%; Score 30; DB 1; Length 362;
CC Best Local Similarity 100.0%; Pred. No. 28;
CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 3 GFYRN 7
CC |||||
CC Db 284 GFYRN 288

```

---

```

Db      284 GFYRN 288
|||||
RESULT 10
SERC_RALSO STANDARD; PRT; 378 AA.
AC Q8Y0Z0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Phosphoserine aminotransferase (EC 2.6.1.52) (PSAT).
GN SERC OR RSC0903 OR RS04512;
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=GMI1000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Denange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Sigulier P., Thébault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
CC Nature 415:497-502(2002).
CC -!- CATALYTIC ACTIVITY: O-phospho-L-serine + 2-oxoglutarate = 3-
CC phosphonooxypyruvate + L-glutamate.
CC -!- COFACTOR: Pyridoxal phosphate (By similarity).
CC -!- PATHWAY: Required both in major phosphorylated pathway of serine
CC biosynthesis and in the biosynthesis of pyridoxine (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to class-V of pyridoxal-phosphate-dependent
CC aminotransferases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL646061; CAD14605.1; -.
CC HAMAP; MF_00160; -.
CC InterPro; IPR000192; Aminotrans V.
CC InterPro; IPR003248; Pser_aminTransf.
CC Pfam; PF00266; aminotran_5; 1.
CC ProDom; PD001544; Pser_amintransf; 1.
CC TIGRFAMs; TIGR01364; serC.1; 1.
CC PROSITE; PS00595; AA_TRANSFER_CLASS_5; FALSE NEG.
CC Serine biosynthesis; Pyridoxine biosynthesis; Transferase;
CC Aminotransferase; Pyridoxal phosphate; Complete proteome.
CC BINDING 214 214 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC SEQUENCE 378 AA; 41860 MW; D4C3356D9098E97B CRC64;
CC
CC Query Match 93.8%; Score 30; DB 1; Length 378;
CC Best Local Similarity 100.0%; Pred. No. 29;
CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 3 GFYRN 7
CC |||||
CC Db 300 GFYRN 304

```

---

```

RESULT 11
PUR2_RHIME STANDARD; PRT; 423 AA.
ID PUR2_RHIME
AC Q9ZRLG;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

DE Phosphoribosylamine--glycine ligase (EC 6.3.4.13) (GARS) (Glycinamide  
DE ribonucleotide synthetase) (Phosphoribosylglycinamide synthetase).  
GN PURD OR R00898 OR SMC00993  
OS Rhizobium meliloti (Sinorhizobium meliloti).  
CC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
CC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.  
OX NCBI\_TaxID=382;  
RN (1)  
RP SEQUENCE FROM N.A.  
PC STRAIN=1021;  
RX MEDLINE=21396507; PubMed=11481430;  
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,  
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,  
RA Pohl T., Portetelle D., Puhler A., Purnelle B., Ramsperger U.,  
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.,  
RA "Analysis of the chromosome sequence of the legume symbiont  
RT Sinorhizobium meliloti strain 1021".  
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
CC -!- CATALYTIC ACTIVITY: ATP + 5-phospho-D-riboseylamine + glycine = ADP  
CC + phosphate + N(1)-(5-phospho-D-riboseyl)glycinamide.  
CC -!- PATHWAY: De novo purine biosynthesis; second step.  
CC -!- SIMILARITY: Belongs to the GARS family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AL591785; CAC45430.1; -;  
DR HAVAP; MF\_00138; -; 1.  
DR InterPro; IPR000115; GARS.  
DR Pfam; PF01071; GARS; 1.  
DR Pfam; PF02842; GARS\_B; 1.  
DR Pfam; PF02843; GARS\_C; 1.  
DR Pfam; PF02844; GARS\_N; 1.  
DR TIGRFAMs; TIGR00877; purD; 1.  
DR PROSITE; PS00184; GARS; 1.  
KW Purine biosynthesis; Ligase; Complete proteome.  
SQ SEQUENCE 423 AA; 44324 MW; 5E65E13B606D204B CRC64;  
  
Query Match 93.8%; Score 30; DB 1; Length 423;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 3 GYFRN 7  
DB 407 GYFRN 411  
|||||  
  
RESULT 12  
SYN MOUSE  
ID SYN MOUSE STANDARD; PRT; 481 AA.  
AC P32E21;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)  
DE (TIPRS).  
DE WARS OR WRS.  
GN Mus musculus (Mouse).  
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN (1)  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RX MEDLINE=95018226; PubMed=7932716;  
RA Pajot B., Sarger C., Bonnet J., Garret M.,  
RT "An alternative splicing modifies the C-terminal end of tryptophanyl-  
tRNA synthetase in murine embryonic stem cells.";

RL J. Mol. Biol. 242:599-603(1994).  
CC -!- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
CC diphosphate + L-tryptophanyl-tRNA(Trp).  
CC -!- SUBUNIT: Homodimer (by similarity).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event-Alternative splicing; Named isoforms=2;  
CC Name=1; Synonyms=Long;  
CC IsoId=P32921-1; Sequences=Displayed;  
CC Name=2; Synonyms=Short;  
CC IsoId=P32921-2; Sequences=VSP 006313;  
CC -!- TISSUE SPECIFICITY: Isoform 2 is widely expressed, isoform 1 is  
CC found only in embryonic stem cells.  
CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.  
CC -!- SIMILARITY: Contains 1 WHEP-TRS domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X69656; CAA49347.1; -;  
DR EMBL; X69657; CAA49348.1; -;  
DR PIR; S50053; S50053.  
DR MGD; MGI:104630; Wars.  
DR InterPro; IPR002305; tRNA-synt\_1b.  
DR InterPro; IPR001412; tRNA-synt\_1.  
DR InterPro; IPR002306; Trp tRNA-synt\_1b.  
DR InterPro; IPR000738; WHEP-TRS.  
DR Pfam; PF00579; tRNA-synt\_1b; 1.  
DR Pfam; PF00458; WHEP-TRS; 1.  
DR PRINTS; PR01039; TRNASYNTHTRP.  
DR TIGRFAMs; TIGR00233; trpS; 1.  
DR PROSITE; PS00178; AA tRNA LIGASE\_I; 1.  
DR PROSITE; PS00762; WHEP TRS; 1.  
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
KW Alternative splicing.  
FT DOMAIN 23 68 WHEP-TRS.  
FT SITE 168 177 "HIGH" REGION.  
FT SITE 353 357 "KMSKS" REGION.  
FT SITE 476 481 Missing (in isoform 2).  
FT VARSPLIC 476 481 /FTId=VSP 006313.  
SQ SEQUENCE 481 AA; 54282 MW; B05A452C08074F52 CRC64;  
  
Query Match 93.8%; Score 30; DB 1; Length 481;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 3 GYFRN 7  
DB 250 GYFRN 254  
|||||  
  
RESULT 13  
TYRO MOUSE  
ID TYRO MOUSE STANDARD; PRT; 533 AA.  
AC P11344;  
DT 01-JUN-1989 (Rel. 11, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Tyrosinase precursor (EC 1.14.18.1) (Monophenol monooxygenase)  
DE (Albino locus protein).  
DE TYR.  
GN Mus musculus (Mouse).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX STRAIN=DEA/2J;  
MEDLINE=88268910; PubMed=3134020;

RA Kwon B.S., Wakulchik M., Haq A.K., Halaban R., Kestler D.;  
RT "Sequence analysis of mouse tyrosinase cDNA and the effect of  
RT melatonin on its gene expression.";  
RL Biochem. Biophys. Res. Commun. 153:1301-1309(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Himalayan;  
RX MEDLINE=89273644; PubMed=2567165;  
RA Kwon B.S., Halaban R., Chintamani C.;  
RT "Molecular basis of mouse Himalayan mutation.";  
RL Biochem. Biophys. Res. Commun. 161:252-260(1989).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89030636; PubMed=3141148;  
RA Mueller G., Ruppert S., Schmid E., Schuetz G.;  
RT "Functional analysis of alternatively spliced tyrosinase gene  
RT transcripts.";  
RL EMBO J. 7:2723-2730(1988).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89193679; PubMed=2494997;  
RA Terao M., Tabé L., Garattini E., Sartori D., Studer M., Mintz B.;  
RT "Isolation and characterization of variant cDNAs encoding mouse  
RT tyrosinase.";  
RL Biochem. Biophys. Res. Commun. 159:848-853(1989).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90212084; PubMed=2517217;  
RA Yamamoto H., Takeuchi S., Kudo T., Makino K., Nakata A., Shinoda T.,  
RA Takeuchi T.;  
RT "Cloning and sequencing of mouse tyrosinase cDNA.";  
RL Jpn. J. Genet. 62:271-274(1987).  
RN [6]  
RP SEQUENCE OF 1-273 FROM N.A.  
RX MEDLINE=90212084; PubMed=2517217;  
RA Yamamoto H., Takeuchi S., Kudo T., Sato C., Takeuchi T.;  
RT "Melanin production in cultured albino melanocytes transfected with  
RT mouse tyrosinase cDNA.";  
RL Jpn. J. Genet. 64:121-135(1989).  
RN [7]  
RP VARIANT ALBINO.  
RC STRAIN=BALB/c;  
RX MEDLINE=90249393; PubMed=2110899;  
RA Shibahara S., Okinaga S., Tomica Y., Takeda A., Yamamoto H., Sato M.,  
RA Takeuchi T.;  
RT "A point mutation in the tyrosinase gene of BALB/c albino mouse  
RT causing the cysteine-->serine substitution at position 85.";  
RL Eur. J. Biochem. 189:455-461(1990).  
RN [8]  
RP VARIANT CHINCHILLA MICE.  
RX MEDLINE=90360993; PubMed=2118105;  
RA Beermann F., Ruppert S., Hummler E., Bosch P.X., Mueller G.,  
RA Ruether U., Schuetz G.;  
RT "Rescue of the albino phenotype by introduction of a functional  
RT tyrosinase gene into mice.";  
RL EMBO J. 9:2819-2826(1990).  
CC -!- FUNCTION: THIS IS A COPPER-CONTAINING OXIDASE THAT FUNCTIONS IN  
CC THE FORMATION OF PIGMENTS SUCH AS MELANINS AND OTHER POLYPHENOLIC  
CC COMPOUNDS. CATALYZES THE RATE-LIMITING CONVERSIONS OF TYROSINE TO  
CC DOPA, DOPA TO DOPA-QUINONE AND POSSIBLY 5,6-DIHYDROXYINDOLE TO  
CC INDOLE-5,6 QUINONE.  
CC -!- CATALYTIC ACTIVITY: L-tyrosine + L-DOPA + O(2) = L-DOPA +  
CC DOPAquinone + H(2)O.  
CC -!- COFACTOR: Binds 2 copper ions per subunit.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Melanosomal.  
CC -!- DISEASE: DEFECTS IN TYR RESULT IN VARIOUS FORMS OF ALBINISM.  
CC HIMALAYAN STRAIN TYROSINASE IS TEMPERATURE-SENSITIVE.  
CC -!- SIMILARITY: Belongs to the tyrosinase family.  
CC -!- CAUTION: REF 4 SEQUENCE WAS INCORRECT DUE TO A DELETION OF EXON 3.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D00440; BAA00341.1; -  
DR EMBL; M20234; AAA40516.1; -  
DR EMBL; M26729; AAA37806.1; -  
DR EMBL; X12782; CAA31273.1; -  
DR EMBL; M24560; AAA40517.1; -  
DR EMBL; D00131; BAA00079.1; -  
DR EMBL; X51743; CAA36033.1; -  
DR EMBL; D00439; BAA00340.1; -  
DR PIR; A27711; YRMSCS.  
DR MGD; MGI:98980; Tyr.  
DR InterPro; IPR008922; Di-copper\_centre.  
DR InterPro; IPR002227; Tyrosinase.  
DR Pfam; PF00264; tyrosinase; 1.  
DR PRINTS; PRO0092; TYROSINASE.  
DR PROSITE; PS00497; TYROSINASE 1; 1.  
DR PROSITE; PS00498; TYROSINASE 2; 1.  
KW Oxidoreductase; Monooxygenase; Copper; Glycoprotein; Signal;  
KW Transmembrane; Melanin biosynthesis; Disease mutation; Albinism.  
FT SIGNAL 1 18  
FT CHAIN 19 533  
FT DOMAIN 19 476  
FT TRANSMEM 477 497  
FT DOMAIN 498 533  
FT METAL 180 180  
FT METAL 202 202  
FT METAL 211 211  
FT METAL 363 363  
FT METAL 367 367  
FT METAL 390 390  
FT DOMAIN 503 508  
FT CARBOHYD 86 86  
FT CARBOHYD 111 111  
FT CARBOHYD 161 161  
FT CARBOHYD 230 230  
FT CARBOHYD 337 337  
FT CARBOHYD 371 371  
FT VARIANT 103 103  
FT VARIANT 420 420  
FT VARIANT 482 482  
FT CONFLICT 40 40  
FT CONFLICT 197 197  
FT CONFLICT 264 264  
FT CONFLICT 346 346  
FT CONFLICT 348 356  
FT CONFLICT 357 403  
FT CONFLICT 471 495  
FT CONFLICT 496 533  
SQ SEQUENCE 533 AA; 60648 MW; 4B711312DDB6F7D1 CRC64;  
Query Match 93.8%; Score 30; DB 1; Length 533;  
Best Local Similarity 100.0%; Pred. No. 41;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 3 GFYRN 7  
Db 458 GFYRN 462  
RESULT 14  
YOTS\_CABEL STANDARD; PRT; 824 AA.  
ID YOTS\_CABEL  
AC P34651;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Hypothetical protein ZK632.5 in chromosome III.  
GN ZK632.5.

OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX MEDLINE=94150718; PubMed=7906398;  
RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M., Cooper J., Coulson A.,  
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,  
RA Fullon L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,  
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,  
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,  
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownteen R.,  
RA Sims M., Smalton M., Smith A., Smith M., Sonhammer E., Staden R.,  
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,  
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,  
RA Wohldman P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans";  
RL Nature 368:32-38(1994).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Z22181; CA80182.1; -;  
DR PIR; S40937; S40937.  
KW Hypothetical protein.  
SQ SEQUENCE 824 AA; 95726 MW; 05074058F5B73919 CRC64;  
Query Match 93.8%; Score 30; DB 1; Length 824;  
Best Local Similarity 100.0%; Pred.No. 64;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 GFYRN 7  
DB 206 GFYRN 210  
RESULT 15  
XPC\_MOUSE  
ID\_XPC\_MOUSE STANDARD; PRT; 900 AA.  
AC P51612; P54732;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE DNA-repair protein complementing XP-C cells homolog (Xeroderma  
DE pigmentosum group C complementing protein homolog) (p125).  
GN XPC.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96184849; PubMed=8604333;  
RA Li L., Peterson C., Legerski R.;  
RT "Sequence of the mouse XPC cDNA and genomic structure of the human  
RT XPC gene";  
RL Nucleic Acids Res. 24:1026-1028(1996).  
RN [2]  
RP SEQUENCE OF 28-587 FROM N.A.  
RC STRAIN=129/Sv;  
RX MEDLINE=95405469; PubMed=7675084;  
RA Sands A.T., Abuin A., Sanchez A., Conti C.J., Bradley A.;  
RT "High susceptibility to ultraviolet-induced carcinogenesis in mice

RT lacking XPC";  
RL Nature 377:162-165(1995).  
CC -1- FUNCTION: Involved in DNA excision repair. May play a part in DNA  
CC damage recognition and/or in altering chromatin structure to  
CC allow access by damage-processing enzymes.  
CC -1- SUBUNIT: HETERODIMER OF A 125 kDa SUBUNIT (P125) AND OF A  
CC 58 kDa SUBUNIT (P58).  
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
CC -1- SIMILARITY: Belongs to the XPC family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U27398; AAC52500.1; -;  
DR EMBL; U40005; AAA82720.1; -;  
DR PIR; S70630; S70630.  
DR MGD; MGI:103557; Xpc.  
DR GO; GO:0006289; P:nucleotide-excision repair; IMP.  
DR InterPro; IPR004583; Rad4.  
DR Pfam; PF03835; Rad4; 1.  
DR TIGRFAMs; TIGR00605; rad4; 1.  
KW DNA repair; DNA-binding; Nuclear protein.  
FT CONFLICT 28 30 AVN -> CSD (IN REF. 2).  
FT CONFLICT 53 53 S -> L (IN REF. 2).  
FT CONFLICT 67 67 L -> F (IN REF. 2).  
FT CONFLICT 70 70 L -> S (IN REF. 2).  
FT CONFLICT 134 135 RG -> TP (IN REF. 2).  
FT CONFLICT 165 170 EQVNM -> GVHEDT (IN REF. 2).  
FT CONFLICT 181 181 S -> N (IN REF. 2).  
FT CONFLICT 187 187 S -> N (IN REF. 2).  
FT CONFLICT 190 190 R -> S (IN REF. 2).  
FT CONFLICT 192 192 P -> L (IN REF. 2).  
FT CONFLICT 342 345 GSKA -> AKP (IN REF. 2).  
FT CONFLICT 367 367 R -> S (IN REF. 2).  
FT CONFLICT 428 428 R -> C (IN REF. 2).  
FT CONFLICT 467 467 C -> S (IN REF. 2).  
FT CONFLICT 584 584 X -> E (IN REF. 2).  
SQ SEQUENCE 900 AA; 100873 MW; 95737FCB36DC15DD CRC64;  
Query Match 93.8%; Score 30; DB 1; Length 900;  
Best Local Similarity 100.0%; Pred.No. 70;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 GFYRN 7  
DB 182 GFYRN 186  
Search completed: September 24, 2004, 07:34:58  
Job time : 4.43158 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 07:28:32 ; Search time 7.29474 Seconds  
(without alignments)  
302.770 Million cell updates/sec

Title: US-09-498-556C-79

Perfect score: 32

Sequence: 1 XXGFYRN 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mbc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_todent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	30	93.8	77	Q8X301	Q8X301 escherichia
2	30	93.8	88	Q8F1E2	Q8F1E2 escherichia
3	30	93.8	91	Q1E938	Q1E938 ancylostoma
4	30	93.8	102	Q962V8	Q962V8 ancylostoma
5	30	93.8	110	Q9KYV8	Q9KYV8 streptomyce
6	30	93.8	113	Q82H68	Q82H68 streptomyce
7	30	93.8	134	P74345	P74345 synchocyst
8	30	93.8	154	Q8ITP8	Q8ITP8 oesophagost
9	30	93.8	158	Q44490	Q44490 anabaena va
10	30	93.8	172	Q9A2A6	Q9A2A6 caulobacter
11	30	93.8	178	Q9TAJ9	Q9TAJ9 careteria r
12	30	93.8	182	Q7V9J4	Q7V9J4 prochloroco
13	30	93.8	184	Q9VGJ2	Q9VGJ2 drosophila
14	30	93.8	184	Q96723	Q96723 drosophila
15	30	93.8	191	Q8L803	Q8L803 triticum ae
16	30	93.8	216	Q8TNS2	Q8TNS2 methanosarc

17	30	93.8	222	16	Q8DIE9	Q8die9 synchococc
18	30	93.8	223	16	Q8YQ65	Q8yq65 anabaena sp
19	30	93.8	228	2	Q8GMR5	Q8gmrs synchococc
20	30	93.8	230	12	Q8UZB5	Q8uzb5 grapevine f
21	30	93.8	237	17	Q8HPV4	Q8hvp4 halobacteri
22	30	93.8	260	16	Q8NTL4	Q8ntl4 corynebacte
23	30	93.8	264	16	Q8FSU2	Q8fsu2 corynebacte
24	30	93.8	266	17	Q8ZT19	Q8zt19 pyrobaculum
25	30	93.8	302	11	Q8C2V6	Q8czv6 mus musculu
26	30	93.8	325	5	Q9BKK3	Q9bbk3 lucilia cup
27	30	93.8	354	5	Q9VEY0	Q9vey0 drosophila
28	30	93.8	381	5	O01482	O01482 caenorhabdi
29	30	93.8	415	5	Q9U153	Q9u153 leishmania
30	30	93.8	465	10	Q8C8J9	Q8c8j9 arabidopsis
31	30	93.8	470	16	Q8EDG2	Q8edg2 shewanella
32	30	93.8	472	3	Q02868	Q02868 saccharomyc
33	30	93.8	475	3	Q9PEK0	Q9pek0 schizosacch
34	30	93.8	475	11	Q9DC65	Q9dc65 mus musculu
35	30	93.8	475	11	Q8OZY4	Q8ozy4 mus musculu
36	30	93.8	475	16	Q8A6N8	Q8a6n8 bacteroides
37	30	93.8	481	11	Q99J58	Q99j58 mus musculu
38	30	93.8	511	10	Q8GRX1	Q8grx1 arabidopsis
39	30	93.8	533	11	Q91XK0	Q91xk0 mus musculu
40	30	93.8	554	17	Q8Z2L2	Q8z2l2 pyrobaculum
41	30	93.8	596	2	O86468	O86468 rhodobacter
42	30	93.8	600	5	Q9NL27	Q9nl27 ciona intes
43	30	93.8	604	16	Q7YPM4	Q7ypm4 haemophilus
44	30	93.8	611	10	Q8GZT8	Q8gztt8 arabidopsis
45	30	93.8	611	10	Q9LT75	Q9lt75 arabidopsis

#### ALIGNMENTS

RESULT 1  
Q8X301 PRELIMINARY; PRT; 77 AA.  
ID Q8X301  
AC Q8X301; 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Hypothetical protein z1836.  
GN Z1836.  
OS Escherichia coli O157:H7.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=83334;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
RX MEDLINE=21074935; PubMed=11206551;  
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouzis K.,  
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
RA Welch R.A., Blattner F.R.;  
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";  
RL Nature 409:529-533(2001).  
DR EMEL; AE005327; AAC55934.1; --  
DR PIR; B95684; B95684.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 77 AA; 8087 MW; F85D56712A58545E CRC64;

Query Match 93.8%; Score 30; DB 16; Length 77;

Best Local Similarity 100.0%; Pred. No. 54;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 GFVRN 7

Db 15 GFVRN 19

```

RESULT 2
Q8FIE2 ID Q8FIE2 PRELIMINARY; PRT; 88 AA.
AC Q8FIE2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN Cl484.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]_TaxID=217992;
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=2238234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
DR EMBL; AB016759; AAN7953.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 88 AA; 9179 MW; 6709ABA0C8259B8B CRC64;

Query Match 93.8%; Score 30; DB 16; Length 88;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 15 GFYRN 19

RESULT 3
Q16938 ID Q16938 PRELIMINARY; PRT; 91 AA.
AC Q16938;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE Anti-coagulant protein C2 precursor (Fragment).
DE Ancylostoma caninum (Dog hookworm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Ancylostomatidae; Ancylostomatinae; Ancylostoma.
OX NCBI_TaxID=29170;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98298519; PubMed=9634780;
RA Jespers L.S., Messens J.H., De Keyser A., Beckhout D.,
RA van den Brande I., Gansemans Y.G., Lauwereys M.J., Vlasuk G.P.,
RA Stanssens P.E.;
RT "Surface expression and ligand-based selection of cDNAs fused to
filamentous phage gene VI.";
RL Biotechnology 13:378-382(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96312555; PubMed=8700900;
RA Stanssens P., Bergum P.W., Gansemans Y., Jaspers L., Larocche Y.,
RA Huang S., Maki S., Messens J., Lauwereys M., Cappello M., Hotez P.J.,
RA Lasters I., Vlasuk G.P.;
RT "Anticoagulant repertoire of the hookworm Ancylostoma caninum.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:2149-2154(1996).
DR EMBL; U30793; AAC47080.1; -.
DR PDB; 1COU; 13-OCT-99.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR002919; TIL_Cysrich.
DR Pfam; PF01826; TIL; 1.
DR PROSITE; PS01186; EGF_2; 1.
KW Signal.

Query Match 93.8%; Score 30; DB 16; Length 88;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 15 GFYRN 19

RESULT 4
Q962V8 ID Q962V8 PRELIMINARY; PRT; 102 AA.
AC Q962V8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Anticoagulant peptide-1 precursor.
GN ACEAP-1.
OS Ancylostoma ceylanicum.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Ancylostomatidae; Ancylostomatinae; Ancylostoma.
OX NCBI_TaxID=53326;
RN [1]
RP SEQUENCE FROM N.A.
RA Harrison L.M., Cappello M.;
RT "Cloning of the major factor Xa inhibitor (AceAP-1) from Ancylostoma
ceylanicum.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF399710; AAK81733.1; -.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR002919; TIL_Cysrich.
DR Pfam; PF01826; TIL; 1.
DR PROSITE; PS01186; EGF_2; 1.
KW Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 102 ANTICOAGULANT PEPTIDE-1.
SQ SEQUENCE 102 AA; 11834 MW; 315722980EF723E7 CRC64;

Query Match 93.8%; Score 30; DB 5; Length 102;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 71 GFYRN 75

RESULT 5
Q9KYV8 ID Q9KYV8 PRELIMINARY; PRT; 110 AA.
AC Q9KYV8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein SCO3187.
GN SCO3187 OR SCE22.04.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycetaceae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidaigo J., Horisby T., Howarth S.,

```

```
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrall B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939115; CAB90971.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 110 AA; 10831 MW; 0893F0F52B8EA8FA CRC64;

Query Match          93.8%; Score 30; DB 16; Length 110;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 GFYRN 7
DB      51 GFYRN 55

RESULT 6
Q82H68          PRELIMINARY; PRT; 113 AA.
AC Q82H68;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN SAV3678.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,
RA "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AP005036; BAC71390.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 113 AA; 11204 MW; CF49EE26F5236DE CRC64;

Query Match          93.8%; Score 30; DB 16; Length 113;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 GFYRN 7
DB      51 GFYRN 55

RESULT 7
P74345          PRELIMINARY; PRT; 134 AA.
AC P74345;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
```

```
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein slr1628.
GN SLR1628.
OS Synecocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Mayajima T., Hirose M., Sugita M., Sasamoto S., Kimura T.,
RA Hosojima T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synecocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
DR EMBL; D90914; BAA18439.1; -.
DR PIR; S76180; S76180.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 134 AA; 14545 MW; 7E5414B80FB43D10 CRC64;

Query Match          93.8%; Score 30; DB 16; Length 134;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 GFYRN 7
DB      27 GFYRN 31

RESULT 8
Q8ITP8          PRELIMINARY; PRT; 154 AA.
AC Q8ITP8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative trypsin-like inhibitor protein precursor.
GN MCRP.
OS Oesophagostomum dentatum.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Strongyloidea; Chabertiidae; Oesophagostomum.
OX NCBI_TaxID=61180;
RN [1]
RP SEQUENCE FROM N.A.
RA Boag P.R., Ranganathan S., Newton S.E., Gasser R.B.;
RT "Identification of a male-specific nematode protein with two trypsin
RT like inhibitor domains.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF399936; AAN32637.1; -.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR002919; TIL_Cysrich.
DR Pfam; PF01826; TIL; 2.
DR PROSITE; PS01186; EGF_2; 1.
KW Signal.
FT SIGNAL.
FT CHAIN.
SQ SEQUENCE 154 AA; 16564 MW; A7F586E3957DA819 CRC64;

Query Match          93.8%; Score 30; DB 5; Length 154;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 GFYRN 7
DB      64 GFYRN 68

RESULT 9
Q44490          PRELIMINARY; PRT; 158 AA.
ID Q44490
```

AC Q44490;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Orf2.  
 OS Anabaena variabilis  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.  
 OX NCBI\_TaxID=1172;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 29413;  
 RX MEDLINE=96016168; PubMed=7566132;  
 RA Thiel T., Lyons E.M., Erker J.C., Ernst A.;  
 RT "A second nitrogenase in vegetative cells of a heterocyst-forming  
 cyanobacterium.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:9358-9362(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 29413;  
 RA Thiel T., Lyons E.M., Erker J.C.;  
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U49859; AAA93026.1; -;  
 DR InterPro; IPR004952; DUF269.  
 DR Pfam; PF03270; DUF269; 1.  
 DR PRODOM; PD008304; DUF269; 1.  
 SQ SEQUENCE 158 AA; 17791 MW; D157EBF59C36FEBD CRC64;

Query Match 93.8%; Score 30; DB 2; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
 Db 35 GFYRN 39

RESULT 10  
 Q9A2A6 PRELIMINARY; PRT; 172 AA.  
 ID Q9A2A6  
 AC Q9A2A6  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein CC3660.  
 GN CC3660.  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;  
 OC Caulobacteraceae; Caulobacter.  
 OX NCBI\_TaxID=155892;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 13089 / CB15;  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Newman M.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Bisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA DeRocker I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA Peto R.J., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
 RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of Caulobacter crescentus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL; AE006024; AA025622.1; -;  
 DR FIC; B87703; B87703.  
 DR TIGR; CC3660; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 172 AA; 18425 MW; 5DD52E712F406F6A CRC64;

Query Match 93.8%; Score 30; DB 16; Length 172;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7

Db 73 GFYRN 77

RESULT 11  
 Q9TAJ9 PRELIMINARY; PRT; 178 AA.  
 ID Q9TAJ9  
 AC Q9TAJ9  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Orf178.  
 GN ORF178.  
 OS Cafeteria roenbergensis.  
 OG Mitochondrion.  
 OC Eukaryota; Stramenopiles; Bicosoecida; Cafeteriaceae; Cafeteria.  
 OX NCBI\_TaxID=33653;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Burger G.;  
 RT "The mitochondrial genome of Cafeteria roenbergensis.";  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF193903; AAF05787.1; -;  
 DR GO; GO:0005739; C:mitochondrion; IEA.  
 KW Mitochondrion.  
 SQ SEQUENCE 178 AA; 21418 MW; 00F69B2FC8461362 CRC64;

Query Match 93.8%; Score 30; DB 8; Length 178;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
 Db 109 GFYRN 113

RESULT 12  
 Q7V9J4 PRELIMINARY; PRT; 182 AA.  
 ID Q7V9J4  
 AC Q7V9J4  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Predicted metal-dependent protease fused to Zn ribbon domain.  
 GN PRO1839.  
 OS Prochlorococcus marinus.  
 OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;  
 OC Prochlorococcus.  
 OX NCBI\_TaxID=1219;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SARG / CCMP 1375 / SS120;  
 RX MEDLINE=22810154; PubMed=12917486;  
 RA Dufréne A., Salanoubat M., Partensky F., Artiguenave F., Axmann I.M.,  
 RA Barbe V., Duprat S., Galperin M.Y., Koonin E.V., Le Gall F.,  
 RA Makarova K.S., Ostrowski M., Ozta S., Robert C., Rogozin I.B.,  
 RA Scanlan D.J., Tandeau de Marsac N., Weissbach J., Wincker P.,  
 RA Wolf Y.I., Hess W.R.;  
 RT "Genome sequence of the cyanobacterium Prochlorococcus marinus SS120,  
 a nearly minimal oxypototrophic genome.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:10020-10025(2003).  
 DR EMBL; AB017166; AAQ00883.1; -;  
 KW Protease; Complete proteome.  
 SQ SEQUENCE 182 AA; 20809 MW; 3D010CB5E5E0B7F6 CRC64;

Query Match 93.8%; Score 30; DB 16; Length 182;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
 Db 44 GFYRN 48

```

RESULT 13
Q9VGJ2 PRELIMINARY; PRT; 184 AA.
ID Q9VGJ2
AC Q9VGJ2:
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE I-T protein.
GN I-T OR CG14719.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Branton R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng X., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fowler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattai B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mearns C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K.A., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos J., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weissbrock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AE003692; AA954687.1; -.
DR FlyBase; FBgn0025821; I-T.
SQ SEQUENCE 184 AA; 21018 MW; 97861392D5B77069 CRC64;

Query Match 93.8%; Score 30; DB 5; Length 184;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 31 GFYRN 35
|||||

RESULT 14
Q96723 PRELIMINARY; PRT; 184 AA.
ID Q96723
AC Q96723:
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Inhibitor-T protein.
GN I-T OR INHIBITOR-T OR CG14719.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CANTON S;
RX MEDLINE=99037736; PubMed=9821974;
RA Helps N.R., Vergidou C., Gaskell T., Cohen P.T.W.;
RT "Characterisation of a novel Drosophila melanogaster testis specific
RT PPI inhibitor related to mammalian inhibitor-2: identification of the
RT site of interaction with PPI."
RL FEBS Lett. 438:131-136(1998).
DR EMBL; AJ006867; CAA07278.1; -.
DR FlyBase; FBgn0025821; I-T.
FT VARIANT 21 21 A -> T.
FT VARIANT 183 183 D -> G.
SQ SEQUENCE 184 AA; 20988 MW; D697139E7271A7B2 CRC64;

Query Match 93.8%; Score 30; DB 5; Length 184;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 31 GFYRN 35
|||||

RESULT 15
Q9L803 PRELIMINARY; PRT; 191 AA.
ID Q9L803
AC Q9L803:
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative plastid ribosomal protein CL9.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RA Lu Z.-X., Laroche A., Gaudet D.;
RT "Triticum aestivum putative plastid ribosomal protein CL9 gene."
RL Submitted (JUN-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AY123421; AAM92711.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005840; C:ribosome; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR InterPro; IPR009027; I9_N-like.
DR InterPro; IPR000244; Ribosomal L9.
DR Pfam; PF01281; Ribosomal L9 C; 1.
DR Pfam; PF01281; Ribosomal L9 N; 1.
DR TIGRFAMs; TIGR00158; L9; 1.
DR PROSITE; PS00651; RIBOSOMAL_L9; 1.
DR Ribosomal protein.
SQ SEQUENCE 191 AA; 21592 MW; DB6D7E72B812A34 CRC64;

Query Match 93.8%; Score 30; DB 10; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 31 GFYRN 35
|||||

```

Fri Sep 24 08:54:05 2004

us-09-498-556c-79.sep04.rspt

Page 6

Db 67 GFYEN 71

Search completed: September 24, 2004, 07:30:25  
Job time : 10.2947 secs



```
Query Match          93.8%; Score 30; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
Db 1 GFYRN 5

RESULT 2
US-08-486-397-70
; Sequence 70, Application US/08486397
; Patent No. 5866542
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Ganssemans, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 357
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,397
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/269
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal fragment
; US-08-486-397-70

Query Match          93.8%; Score 30; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
Db 1 GFYRN 5

RESULT 4
US-08-461-965-70
; Sequence 70, Application US/08461965
; Patent No. 5872098
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Ganssemans, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
```

STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
CURRENT APPLICATION DATA:  
SOFTWARE: Word Perfect 5.1  
APPLICATION NUMBER: US/08/461,965  
FILING DATE: June 5, 1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 210/243  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 70:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal fragment  
US-08-461-965-70

Query Match 93.8%; Score 30; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0;

QY 3 GFYRN 7  
Db 1 GFYRN 5

RESULT 5  
US-08-634-641-70  
Sequence 70, Application US/08634641  
Patent No. 5955294  
GENERAL INFORMATION:  
APPLICANT: Vlasuk, George P. Vlasuk  
APPLICANT: Stanssens, Patrick Eric Hugo  
APPLICANT: Menssens, Joris Hilda Llieven  
APPLICANT: Lauwereys, Marc Josef  
APPLICANT: Laroche, Yves Rene  
APPLICANT: Jespers, Laurent Stephane  
APPLICANT: Ganssemans, Yannick Georges Jozef  
APPLICANT: Moyle, Matthew  
APPLICANT: Bergum, Peter W.  
TITLE OF INVENTION: NEMATOBE-EXTRACTED ANTICOAGULANT  
TITLE OF INVENTION: PROTEIN  
NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/634,641  
FILING DATE: April 19, 1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486,397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 219/136  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 70:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal fragment  
US-08-634-641-70

Query Match 93.8%; Score 30; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0;

QY 3 GFYRN 7  
Db 1 GFYRN 5

RESULT 6  
US-09-249-471-70  
Sequence 70, Application US/09249471  
Patent No. 6040441  
GENERAL INFORMATION:  
APPLICANT: Vlasuk, George Phillip  
APPLICANT: Stanssens, Patrick Eric Hugo  
APPLICANT: Messens, Joris Hilda Llieven  
APPLICANT: Lauwereys, Marc Josef  
APPLICANT: Laroche, Yves Rene  
APPLICANT: Jespers, Laurent Stephane  
APPLICANT: Ganssemans, Yannick Georges Jozef  
APPLICANT: Moyle, Matthew  
APPLICANT: Bergum, Peter W.  
TITLE OF INVENTION: NEMATOBE-EXTRACTED SERINE PROTEASE  
TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT  
NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage

COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/249,471  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/809,455  
FILING DATE: April 17, 1997  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486,397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 70:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal fragment  
US-09-249-471-70

Query Match 93.8%; Score 30; DB 3; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
Db 1 GFYRN 5

RESULT 7  
US-09-249-472-70  
Sequence 70, Application US/09249472  
Patent No. 6046318  
GENERAL INFORMATION:  
APPLICANT: Vlasuk, George Phillip  
APPLICANT: Stanssens, Patrick Eric Hugo  
APPLICANT: Messens, Joris Hilda Lieven  
APPLICANT: Lauwereys, Marc Josef  
APPLICANT: Larocche, Yves Rene  
APPLICANT: Jespers, Laurent Stephane  
APPLICANT: Ganssemans, Yannick Georges Jozef  
APPLICANT: Moyle, Matthew  
APPLICANT: Bergum, Peter W.  
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT  
NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: Storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/249,472  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/809,455  
FILING DATE: April 17, 1997  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486,397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 70:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal fragment  
US-09-249-472-70

Query Match 93.8%; Score 30; DB 3; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
Db 1 GFYRN 5

RESULT 8  
US-09-249-451-70  
Sequence 70, Application US/09249451  
Patent No. 6087487  
GENERAL INFORMATION:  
APPLICANT: Vlasuk, George Phillip  
APPLICANT: Stanssens, Patrick Eric Hugo  
APPLICANT: Messens, Joris Hilda Lieven  
APPLICANT: Lauwereys, Marc Josef  
APPLICANT: Larocche, Yves Rene  
APPLICANT: Jespers, Laurent Stephane  
APPLICANT: Ganssemans, Yannick Georges Jozef  
APPLICANT: Moyle, Matthew  
APPLICANT: Bergum, Peter W.  
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT  
NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles

STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/249,451  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/809,455  
FILING DATE: April 17, 1997  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486,397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 70:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal fragment  
US-09-249-451-70

Query Match 93.8%; Score 30; DB 3; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 3 GFYRN 7  
Db 1 GFYRN 5

RESULT 9  
US-08-809-455-70  
Sequence 70, Application US/08809455  
Patent No. 6090916  
GENERAL INFORMATION:  
APPLICANT: Vlasuk, George Phillip  
APPLICANT: Stanssens, Patrick Eric Hugo  
APPLICANT: Messens, Joris Hilda Lieven  
APPLICANT: Lauwereys, Marc Josef  
APPLICANT: Laroche, Yves Rene  
APPLICANT: Jespers, Laurent Stephane  
APPLICANT: Ganssemans, Yannick Georges Jozef  
APPLICANT: Moyle, Matthew  
APPLICANT: Bergum, Peter W.  
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT  
TITLE OF INVENTION: PROTEIN  
NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/809,455  
FILING DATE: April 17, 1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486,397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 70:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal fragment  
US-08-809-455-70

Query Match 93.8%; Score 30; DB 3; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 3 GFYRN 7  
Db 1 GFYRN 5

RESULT 10  
US-09-249-461-70  
Sequence 70, Application US/09249461  
Patent No. 6096877  
GENERAL INFORMATION:  
APPLICANT: Vlasuk, George Phillip  
APPLICANT: Stanssens, Patrick Eric Hugo  
APPLICANT: Messens, Joris Hilda Lieven  
APPLICANT: Lauwereys, Marc Josef  
APPLICANT: Laroche, Yves Rene  
APPLICANT: Jespers, Laurent Stephane  
APPLICANT: Ganssemans, Yannick Georges Jozef  
APPLICANT: Moyle, Matthew  
APPLICANT: Bergum, Peter W.  
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT  
TITLE OF INVENTION: PROTEIN  
NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:

us-09-498-556c-79.sep04.ra1

ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/249,461  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/809,455  
FILING DATE: April 17, 1997  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486,397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 70:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal fragment  
US-09-249-461-70

Query Match 93.8%; Score 30; DB 3; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
Db 1 GFYRN 5

## RESULT 11

US-09-249-448-70  
Sequence 70, Application US/09249448  
Patent No. 6121435  
GENERAL INFORMATION:  
APPLICANT: Vlasuk, George Phillip  
APPLICANT: Stanssens, Patrick Eric Hugo  
APPLICANT: Messens, Joris Hilda Llieven  
APPLICANT: Lauwereys, Marc Josef  
APPLICANT: Laroche, Yves Rene  
APPLICANT: Jespers, Laurent Stephane  
APPLICANT: Gansemans, Yannick Georges Jozef  
APPLICANT: Moyle, Matthew  
APPLICANT: Beigum, Peter W.  
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
INHIBITORS AND ANTICOAGULANT

TITLE OF INVENTION: PROTEIN  
NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/249,448  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/809,455  
FILING DATE: April 17, 1997  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486,397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 70:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal fragment  
US-09-249-448-70

Query Match 93.8%; Score 30; DB 3; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
Db 1 GFYRN 5

## RESULT 12

US-09-249-473-70  
Sequence 70, Application US/09249473  
Patent No. 6534629  
GENERAL INFORMATION:  
APPLICANT: Vlasuk, George Phillip  
APPLICANT: Stanssens, Patrick Eric Hugo  
APPLICANT: Messens, Joris Hilda Llieven  
APPLICANT: Lauwereys, Marc Josef  
APPLICANT: Laroche, Yves Rene  
APPLICANT: Jespers, Laurent Stephane  
APPLICANT: Gansemans, Yannick Georges Jozef  
APPLICANT: Moyle, Matthew

APPLICANT: Bergum, Peter W.  
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT  
NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/249,473  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/909,455  
FILING DATE: April 17, 1997  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486,397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 70:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal fragment  
US-09-249-473-70

Query Match 93.8%; Score 30; DB 4; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
Db 1 GFYRN 5

RESULT 13  
US-08-465-380-79  
Sequence 79, Application US/08465380  
Patent No. 5863894  
GENERAL INFORMATION:  
APPLICANT: George P. Vlasuk, Patric H. Stanssens,  
APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,  
APPLICANT: Yves R. Laroche, Laurent S. Jespers,  
APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,  
APPLICANT: Peter W. Bergum

TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT  
TITLE OF INVENTION: PROTEIN  
NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/465,380  
FILING DATE: June 5, 1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 213/268  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 79:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal fragment  
FEATURE:  
OTHER INFORMATION: Xaa in locations 1 and 2  
OTHER INFORMATION: is an amino acid, provided  
OTHER INFORMATION: that at least one Xaa is Glu or  
OTHER INFORMATION: Asp.  
US-08-465-380-79

Query Match 93.8%; Score 30; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
Db 3 GFYRN 7

RESULT 14  
US-08-486-397-79  
Sequence 79, Application US/08486397  
Patent No. 586542  
GENERAL INFORMATION:  
APPLICANT: George P. Vlasuk, Patric H. Stanssens,  
APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,  
APPLICANT: Yves R. Laroche, Laurent S. Jespers,  
APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,  
APPLICANT: Peter W. Bergum  
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT  
NUMBER OF SEQUENCES: 357  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles

```

/ STATE: California
/ COUNTRY: U.S.A.
/ ZIP: 90071
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
/ MEDIUM TYPE: storage
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: IBM P.C. DOS 5.0
/ SOFTWARE: Word Perfect 5.1
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/486,397
/ FILING DATE: June 5, 1995
/ CLASSIFICATION: 530
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/326,110
/ FILING DATE: October 18, 1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BIGGS, SUZANNE L.
/ REGISTRATION NUMBER: 30,158
/ REFERENCE/DOCKET NUMBER: 213/270
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (213) 489-1600
/ TELEFAX: (213) 955-0440
/ TELEX: 67-3510
/ INFORMATION FOR SEQ ID NO: 79:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 7 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ FRAGMENT TYPE: internal fragment
/ FEATURE:
/ OTHER INFORMATION: Xaa in locations 1 and 2
/ OTHER INFORMATION: is an amino acid, provided
/ OTHER INFORMATION: that at least one Xaa is Glu or
/ OTHER INFORMATION: Asp.
/ US-08-486-397-79

Query Match 93.8%; Score 30; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
|||||
DB 3 GFYRN 7

Search completed: September 24, 2004, 07:50:01
Job time: 5.34737 secs

```

```

RESULT 15
US-08-486-399-79
; Sequence 79, Application US/08496399
; Patent No. 5865543
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Coris H.I. Mensens, Marc J. Leuwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Ganssemans, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 07:35:06 ; Search time 60.2737 Seconds  
(without alignments)  
37.345 Million cell updates/sec

Title: US-09-498-556C-79

Perfect score: 32

Sequence: 1 XXGYRN 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 321558718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	93.8	5	10	US-09-498-272-70
2	30	93.8	7	10	US-09-498-272-79
3	30	93.8	78	10	US-09-498-272-47
4	30	93.8	82	10	US-09-498-272-25
5	30	93.8	82	10	US-09-498-272-49
6	30	93.8	83	10	US-09-498-272-52
7	30	93.8	84	10	US-09-498-272-50
8	30	93.8	84	10	US-09-498-272-51
9	30	93.8	84	10	US-09-498-272-59
10	30	93.8	87	12	US-10-424-599-249825
11	30	93.8	89	10	US-09-498-272-46
12	30	93.8	91	10	US-09-498-272-128
13	30	93.8	92	12	US-10-424-599-263776
14	30	93.8	102	10	US-09-498-272-27
15	30	93.8	108	10	US-09-498-272-24

16	30	93.8	113	14	US-10-156-761-11215
17	30	93.8	138	9	US-09-764-877-1934
18	30	93.8	138	15	US-10-242-515-1934
19	30	93.8	149	16	US-10-437-963-130069
20	30	93.8	161	10	US-09-498-272-65
21	30	93.8	162	10	US-09-498-272-63
22	30	93.8	162	10	US-09-498-272-64
23	30	93.8	171	10	US-09-498-272-62
24	30	93.8	185	15	US-10-629-951-35
25	30	93.8	192	16	US-10-437-963-147735
26	30	93.8	240	12	US-10-424-599-159958
27	30	93.8	243	16	US-10-767-701-43449
28	30	93.8	252	12	US-10-425-114-36701
29	30	93.8	260	9	US-09-738-626-3828
30	30	93.8	268	9	US-09-864-866-43
31	30	93.8	354	15	US-10-369-493-9103
32	30	93.8	359	15	US-10-369-493-9498
33	30	93.8	359	15	US-10-369-493-15415
34	30	93.8	361	12	US-10-282-122A-78000
35	30	93.8	361	15	US-10-369-493-15783
36	30	93.8	361	15	US-10-369-493-16164
37	30	93.8	362	15	US-10-369-493-17731
38	30	93.8	382	16	US-10-437-963-120180
39	30	93.8	475	14	US-10-205-219-65
40	30	93.8	533	12	US-09-961-656-5
41	30	93.8	533	14	US-10-207-655-75
42	30	93.8	623	12	US-10-424-599-238442
43	30	93.8	795	12	US-10-398-471-1
44	30	93.8	795	12	US-10-398-471-3
45	30	93.8	795	12	US-10-398-471-13

#### ALIGNMENTS

#### RESULT 1

US-09-498-272-70

; Sequence 70, Application US/09498272

; Publication No. US20030113890A1

; GENERAL INFORMATION:

APPLICANT: Vlasuk, George Phillip

Stanssens, Patrick Eric Hugo

Messens, Joris Hilda Lieven

Lauwerijs, Marc Josef

Laroche, Yves Rene

Jespers, Laurent Stephane

Ganseman, Yannick Georges Jozef

Moyle, Matthew

Bergum, Peter W.

TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE

INHIBITORS AND ANTICOAGULANT

PROTEIN

NUMBER OF SEQUENCES: 356

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street

Suite 4700

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/498,272

FILING DATE: 04-Feb-2000

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/13231

FILING DATE: October 17, 1995

```

/ APPLICATION NUMBER: 08/486,399
/ FILING DATE: June 5, 1995
/ APPLICATION NUMBER: 08/486,399
/ FILING DATE: June 5, 1995
/ APPLICATION NUMBER: 08/486,397
/ FILING DATE: June 5, 1995
/ APPLICATION NUMBER: 08/465,380
/ FILING DATE: June 5, 1995
/ APPLICATION NUMBER: 08/461,965
/ FILING DATE: June 5, 1995
/ APPLICATION NUMBER: 08/326,110
/ FILING DATE: October 18, 1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BIGGS, SUZANNE L.
/ REGISTRATION NUMBER: 30,158
/ REFERENCE/DOCKET NUMBER: 216/270
/ TELEPHONE: (213) 489-1600
/ TELEFAX: (213) 955-0440
/ TELEX: 67-3510
/ INFORMATION FOR SEQ ID NO: 70:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 5 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ FRAGMENT TYPE: internal fragment
/ SEQUENCE DESCRIPTION: SEQ ID NO: 70:
US-09-498-272-70
Query Match 93.8%; Score 30; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.2e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GYRN 7
DB 1 GYRN 5

RESULT 2
US-09-498-272-79
/ Sequence 79, Application US/09498272
/ Publication No. US20030113890A1
/ GENERAL INFORMATION:
/ APPLICANT: Vlasuk, George Phillip
/ Stanssens, Patrick Eric Hugo
/ Messens, Joris Hilda Lieven
/ Lauwereys, Marc Josef
/ Laroche, Yves Rene
/ Jespers, Laurent Stephane
/ Ganssemans, Yannick Georges Jozef
/ Moyle, Matthew
/ Bergum, Peter W.
/ TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
/ INHIBITORS AND ANTICOAGULANT
/ PROTEIN
/ NUMBER OF SEQUENCES: 356
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Lyon & Lyon
/ STREET: 633 West Fifth Street
/ CITY: Suite 4700
/ CITY: Los Angeles
/ STATE: California
/ COUNTRY: U.S.A.
/ ZIP: 90071
/ MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: IBM P.C. DOS 5.0
/ SOFTWARE: Word Perfect 5.1
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/498,272
/ FILING DATE: 04-Feb-2000
/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: PCT/US95/13231
/ FILING DATE: October 17, 1995
/ APPLICATION NUMBER: 08/486,399
/ FILING DATE: June 5, 1995
/ APPLICATION NUMBER: 08/486,397
/ FILING DATE: June 5, 1995
/ APPLICATION NUMBER: 08/465,380
/ FILING DATE: June 5, 1995
/ APPLICATION NUMBER: 08/461,965
/ FILING DATE: June 5, 1995
/ APPLICATION NUMBER: 08/326,110
/ FILING DATE: October 18, 1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BIGGS, SUZANNE L.
/ REGISTRATION NUMBER: 30,158
/ REFERENCE/DOCKET NUMBER: 216/270
/ TELEPHONE: (213) 489-1600
/ TELEFAX: (213) 955-0440
/ TELEX: 67-3510
/ INFORMATION FOR SEQ ID NO: 79:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 7 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ FRAGMENT TYPE: internal fragment
/ FEATURE:
/ OTHER INFORMATION: Xaa in locations 1 and 2
/ is an amino acid, provided
/ that at least one Xaa is Glu or
/ Asp.
/ SEQUENCE DESCRIPTION: SEQ ID NO: 79:
US-09-498-272-79
Query Match 93.8%; Score 30; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.2e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GYRN 7
DB 3 GYRN 7

RESULT 3
US-09-498-272-47
/ Sequence 47, Application US/09498272
/ Publication No. US20030113890A1
/ GENERAL INFORMATION:
/ APPLICANT: Vlasuk, George Phillip
/ Stanssens, Patrick Eric Hugo
/ Messens, Joris Hilda Lieven
/ Lauwereys, Marc Josef
/ Laroche, Yves Rene
/ Jespers, Laurent Stephane
/ Ganssemans, Yannick Georges Jozef
/ Moyle, Matthew
/ Bergum, Peter W.
/ TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
/ INHIBITORS AND ANTICOAGULANT
/ PROTEIN
/ NUMBER OF SEQUENCES: 356
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Lyon & Lyon
/ STREET: 633 West Fifth Street
/ CITY: Suite 4700
/ CITY: Los Angeles
/ STATE: California
/ COUNTRY: U.S.A.
/ ZIP: 90071
/ MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
/ COMPUTER READABLE FORM:
/ storage
```

COMPUTER: IBM Compatible  
 OPERATING SYSTEM: IBM P.C. DOS 5.0  
 SOFTWARE: Word Perfect 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/498,272  
 FILING DATE: 04-Feb-2000  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US95/13231  
 FILING DATE: October 17, 1995  
 APPLICATION NUMBER: 08/486,399  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/486,397  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/465,380  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/461,965  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/326,110  
 FILING DATE: October 18, 1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BIGGS, SUZANNE L.  
 REGISTRATION NUMBER: 30,158  
 REFERENCE/DOCKET NUMBER: 216/270  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (213) 489-1600  
 TELEFAX: (213) 955-0440  
 TELEX: 67-3510  
 INFORMATION FOR SEQ ID NO: 47:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 78 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 ORIGINAL SOURCE:  
 ORGANISM: Ancylostoma caninum  
 SEQUENCE DESCRIPTION: SEQ ID NO: 47:  
 US-09-498-272-47

Query Match 93.8%; Score 30; DB 10; Length 78;  
 Best Local Similarity 100.0%; Pred. No. 55;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GYRN 7  
 DB 54 GYRN 58

RESULT 4  
 US-09-498-272-25  
 Sequence 25, Application US/09498272  
 Publication No. US20030113890A1  
 GENERAL INFORMATION:  
 APPLICANT: Vlasuk, George Phillip  
 Stanssens, Patrick Eric Hugo  
 Messens, Joris Hilda Lieven  
 Lauwereys, Marc Josef  
 Laroche, Yves Rene  
 Jespers, Laurent Stephane  
 Ganssemans, Yannick Georges Jozef  
 Moyle, Matthew  
 Bergum, Peter W.  
 TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
 INHIBITORS AND ANTICOAGULANT  
 PROTEIN  
 NUMBER OF SEQUENCES: 356  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Lyon & Lyon  
 STREET: 633 West Fifth Street  
 Suite 4700  
 CITY: Los Angeles  
 STATE: California  
 COUNTRY: U.S.A.  
 ZIP: 90071

COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: IBM P.C. DOS 5.0  
 SOFTWARE: Word Perfect 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/498,272  
 FILING DATE: 04-Feb-2000  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US95/13231  
 FILING DATE: October 17, 1995  
 APPLICATION NUMBER: 08/486,399  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/486,397  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/465,380  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/461,965  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/326,110  
 FILING DATE: October 18, 1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BIGGS, SUZANNE L.  
 REGISTRATION NUMBER: 30,158  
 REFERENCE/DOCKET NUMBER: 216/270  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (213) 489-1600  
 TELEFAX: (213) 955-0440  
 TELEX: 67-3510  
 INFORMATION FOR SEQ ID NO: 25:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 82 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 ORIGINAL SOURCE:  
 ORGANISM: Ancylostoma ceylanicum  
 SEQUENCE DESCRIPTION: SEQ ID NO: 25:  
 US-09-498-272-25

Query Match 93.8%; Score 30; DB 10; Length 82;  
 Best Local Similarity 100.0%; Pred. No. 58;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GYRN 7  
 DB 51 GYRN 55

RESULT 5  
 US-09-498-272-49  
 Sequence 49, Application US/09498272  
 Publication No. US20030113890A1  
 GENERAL INFORMATION:  
 APPLICANT: Vlasuk, George Phillip  
 Stanssens, Patrick Eric Hugo  
 Messens, Joris Hilda Lieven  
 Lauwereys, Marc Josef  
 Laroche, Yves Rene  
 Jespers, Laurent Stephane  
 Ganssemans, Yannick Georges Jozef  
 Moyle, Matthew  
 Bergum, Peter W.  
 TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
 INHIBITORS AND ANTICOAGULANT  
 PROTEIN  
 NUMBER OF SEQUENCES: 356  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Lyon & Lyon  
 STREET: 633 West Fifth Street  
 Suite 4700  
 CITY: Los Angeles

STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/498,272  
FILING DATE: 04-Feb-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486,397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 82 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Ancylostoma ceylanicum  
SEQUENCE DESCRIPTION: SEQ ID NO: 49:  
US-09-498-272-49

Query Match 93.8%; Score 30; DB 10; Length 82;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
Db 51 GFYRN 55

## RESULT 6

US-09-498-272-52  
Sequence 52, Application US/09498272  
Publication No. US20030113890A1

## GENERAL INFORMATION:

APPLICANT: Vlasuk, George Phillip  
Stanssens, Patrick Eric Hugo  
Messens, Joris Hilda Lieven  
Lauwereys, Marc Josef  
Laroche, Yves Rene  
Jespers, Laurent Stephane  
Gansemans, Yannick Georges Jozef  
Moyle, Matthew  
Bergum, Peter W.

TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
INHIBITORS AND ANTICOAGULANT  
PROTEIN

NUMBER OF SEQUENCES: 356  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street  
CITY: Suite 4700  
STATE: Los Angeles  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/498,272  
FILING DATE: 04-Feb-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13231  
FILING DATE: October 17, 1995  
APPLICATION NUMBER: 08/486,399  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/486,397  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/465,380  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/461,965  
FILING DATE: June 5, 1995  
APPLICATION NUMBER: 08/326,110  
FILING DATE: October 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 52:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 83 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Ancylostoma duodenale  
SEQUENCE DESCRIPTION: SEQ ID NO: 52:  
US-09-498-272-52

Query Match 93.8%; Score 30; DB 10; Length 83;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7  
Db 55 GFYRN 59

## RESULT 7

US-09-498-272-50  
Sequence 50, Application US/09498272  
Publication No. US20030113890A1

## GENERAL INFORMATION:

APPLICANT: Vlasuk, George Phillip  
Stanssens, Patrick Eric Hugo  
Messens, Joris Hilda Lieven  
Lauwereys, Marc Josef  
Laroche, Yves Rene  
Jespers, Laurent Stephane  
Gansemans, Yannick Georges Jozef  
Moyle, Matthew  
Bergum, Peter W.

TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
INHIBITORS AND ANTICOAGULANT  
PROTEIN

Query Match	93.8%	Score 30;	DB 10;	Length 84;
Best Local Similarity	100.0%;	Pred. No. 59;		
Match Conservative	0.	Mismatches	0.	Indels
Mismatch	0.	Gaps	0.	

RESULT 9  
US-09-498-272-59  
; Sequence 59, Application US/09498272  
; Publication No. US2003013890A1  
; GENERAL INFORMATION:  
; APPLICANT: Vlasuk, George Phillip  
; Starszens, Patrick Eric Hugo  
; Messens, Joris Hilda Hieven  
; Lauwereys, Marc Josef  
; Laroche, Yves Rene  
; Jespers, Laurent Stephane

Query Match	93.8%	Score 30;	DB 10;	Length 84;
Best Local Similarity	100.0%;	Pred. No. 59;		
Matches	5;	Conservative	0;	Mismatches
			0;	Indels
			0;	Gaps
			0;	Gaps

```

RESULT 8
US-09-498-272-51
; Sequence 51, Application US/09498272
; Publication NO. US20030113890A1
;
GENERAL INFORMATION:
;
APPLICANT: Vlasuk, George Phillip
;
; Stanssens, Patrick Eric Hugo
; Messens, Joris Hilda Lieveen
; Lauwereys, Marc Josef
; Laroche, Yves Rene
; Jespers, Laurent Stephane
; Ganssemaus, Yannick Georges Jozef
; Moyle, Matthew
; Barum, Peter W.
;

```

```

; Gansemans, Yannick Georges Jozef
; Moyle, Matthew
; Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; INHIBITORS AND ANTICOAGULANT
;
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; SUITE: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA: US/09/498, 272
; FILING DATE: 04-Feb-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
;
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
;
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; SEQUENCE DESCRIPTION: SEQ ID NO: 59:
US-09-498-272-59

```

```

Query Match          93.8%; Score 30; DB 10; Length 84;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      3 GFYRN 7
      |||||
Db      55 GFYRN 59

```

```

RESULT 10
US-10-424-599-249825
; Sequence 249825, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua

```

```

; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 249825
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_67620C.1.pep
US-10-424-599-249825

```

```

Query Match          93.8%; Score 30; DB 12; Length 87;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      3 GFYRN 7
      |||||
Db      63 GFYRN 67

```

```

RESULT 11
US-09-498-272-48
; Sequence 48, Application US/09498272
; Publication No. US20030113890A1
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George Phillip
; Stanssens, Patrick Eric Hugo
; Messens, Joris Hilda Lieven
; Lauwereys, Marc Josef
; Laroche, Yves Rene
; Jespers, Laurent Stephane
; Gansemans, Yannick Georges Jozef
; Moyle, Matthew
; Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; INHIBITORS AND ANTICOAGULANT
; PROTEIN

```

```

; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; SUITE: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA: US/09/498, 272
; FILING DATE: 04-Feb-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
;
; ATTORNEY/AGENT INFORMATION:

```

NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 89 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Ancylostoma ceylanicum  
SEQUENCE DESCRIPTION: SEQ ID NO: 48:  
US-09-498-272-48

Query Match 93.8%; Score 30; DB 10; Length 89;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GYRN 7  
Db 61 GYRN 65

RESULT 12  
US-09-498-272-128  
; Sequence 128, Application US/09498272  
; Publication No. US20030113890A1  
; GENERAL INFORMATION:  
; APPLICANT: Vlasuk, George Phillip  
; Stanssens, Patrick Eric Hugo  
; Messens, Joris Hilda Lieven  
; Lauwereys, Marc Josef  
; Jaspers, Laurent Stephane  
; Ganssens, Yannick Georges Jozef  
; Moyle, Matthew  
; Bergum, Peter W.  
; TITLE OF INVENTION: NEUTRODE-EXTRACTED SERINE PROTEASE  
; INHIBITORS AND ANTICOAGULANT  
; PROTEIN  
; NUMBER OF SEQUENCES: 356  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; CITY: Suite 4700  
; STATE: Los Angeles  
; COUNTRY: California  
; ZIP: 90071  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/498,272  
; FILING DATE: 04-Feb-2000  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/13231  
; FILING DATE: October 17, 1995  
; APPLICATION NUMBER: 08/486,399  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/486,397  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/465,380  
; FILING DATE: June 5, 1995  
; APPLICATION NUMBER: 08/461,965  
; FILING DATE: June 5, 1995

NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 216/270  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 128:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 91 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Ancylostoma caninum  
SEQUENCE DESCRIPTION: SEQ ID NO: 128:  
US-09-498-272-128

Query Match 93.8%; Score 30; DB 10; Length 91;  
Best Local Similarity 100.0%; Pred. No. 64;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GYRN 7  
Db 62 GYRN 66

RESULT 13  
US-10-424-599-263776  
; Sequence 263776, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 263776  
; LENGTH: 92  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_80209C.1.pcp  
US-10-424-599-263776

Query Match 93.8%; Score 30; DB 12; Length 92;  
Best Local Similarity 100.0%; Pred. No. 65;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GYRN 7  
Db 19 GYRN 23

RESULT 14  
US-09-498-272-27  
; Sequence 27, Application US/09498272  
; Publication No. US20030113890A1  
; GENERAL INFORMATION:  
; APPLICANT: Vlasuk, George Phillip  
; Stanssens, Patrick Eric Hugo  
; Messens, Joris Hilda Lieven  
; Lauwereys, Marc Josef  
; Laroche, Yves Rene  
; Jespers, Laurent Stephane

/ Gansemans, Yannick Georges Jozef  
/ Moyle, Matthew  
/ Bergum, Peter W.  
/ TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
/ INHIBITORS AND ANTICOAGULANT  
/ PROTEIN  
/ NUMBER OF SEQUENCES: 356  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: Lyon & Lyon  
/ STREET: 633 West Fifth Street  
/ Suite 4700  
/ CITY: Los Angeles  
/ STATE: California  
/ COUNTRY: U.S.A.  
/ ZIP: 90071  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
/ storage  
/ COMPUTER: IBM Compatible  
/ OPERATING SYSTEM: IBM P.C. DOS 5.0  
/ SOFTWARE: Word Perfect 5.1  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/09/498,272  
/ FILING DATE: 04-Feb-2000  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: PCT/US95/13231  
/ FILING DATE: October 17, 1995  
/ APPLICATION NUMBER: 08/486,399  
/ FILING DATE: June 5, 1995  
/ APPLICATION NUMBER: 08/486,397  
/ FILING DATE: June 5, 1995  
/ APPLICATION NUMBER: 08/465,380  
/ FILING DATE: June 5, 1995  
/ APPLICATION NUMBER: 08/326,110  
/ FILING DATE: October 18, 1994  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: BIGGS, SUZANNE L.  
/ REGISTRATION NUMBER: 30,158  
/ REFERENCE/DOCKET NUMBER: 216/270  
/ TELEPHONE: (213) 489-1600  
/ TELEFAX: (213) 955-0440  
/ TELEX: 67-3510  
/ INFORMATION FOR SEQ ID NO: 27:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 102 amino acids  
/ TYPE: amino acid  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: peptide  
/ ORIGINAL SOURCE:  
/ ORGANISM: Ancylostoma duodenale  
/ SEQUENCE DESCRIPTION: SEQ ID NO: 27:  
US-09-498-272-27

Query Match 93.8%; Score 30; DB 10; Length 102;  
Best Local Similarity 100.0%; Pred. No. 72;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7

Db 74 GFYRN 78

RESULT 15

US-09-498-272-24  
/ Sequence 24, Application US/09498272  
/ Publication No. US20030113890A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Vlasuk, George Phillip  
/ Stanssens, Patrick Eric Hugo  
/ Messens, Joris Hilda Lieven

/ Lauwereys, Marc Josef  
/ Laroche, Yves Rene  
/ Jespers, Laurent Stephane  
/ Gansemans, Yannick Georges Jozef  
/ Moyle, Matthew  
/ Bergum, Peter W.  
/ TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
/ INHIBITORS AND ANTICOAGULANT  
/ PROTEIN  
/ NUMBER OF SEQUENCES: 356  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: Lyon & Lyon  
/ STREET: 633 West Fifth Street  
/ Suite 4700  
/ CITY: Los Angeles  
/ STATE: California  
/ COUNTRY: U.S.A.  
/ ZIP: 90071  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
/ storage  
/ COMPUTER: IBM Compatible  
/ OPERATING SYSTEM: IBM P.C. DOS 5.0  
/ SOFTWARE: Word Perfect 5.1  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/09/498,272  
/ FILING DATE: 04-Feb-2000  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: PCT/US95/13231  
/ FILING DATE: October 17, 1995  
/ APPLICATION NUMBER: 08/486,399  
/ FILING DATE: June 5, 1995  
/ APPLICATION NUMBER: 08/486,397  
/ FILING DATE: June 5, 1995  
/ APPLICATION NUMBER: 08/465,380  
/ FILING DATE: June 5, 1995  
/ APPLICATION NUMBER: 08/461,965  
/ FILING DATE: June 5, 1995  
/ APPLICATION NUMBER: 08/326,110  
/ FILING DATE: October 18, 1994  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: BIGGS, SUZANNE L.  
/ REGISTRATION NUMBER: 30,158  
/ REFERENCE/DOCKET NUMBER: 216/270  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (213) 489-1600  
/ TELEFAX: (213) 955-0440  
/ TELEX: 67-3510  
/ INFORMATION FOR SEQ ID NO: 24:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 108 amino acids  
/ TYPE: amino acid  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: peptide  
/ ORIGINAL SOURCE:  
/ ORGANISM: Ancylostoma ceylanicum  
/ SEQUENCE DESCRIPTION: SEQ ID NO: 24:  
US-09-498-272-24

Query Match 93.8%; Score 30; DB 10; Length 108;  
Best Local Similarity 100.0%; Pred. No. 76;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7

Db 80 GFYRN 84

Search completed: September 24, 2004, 08:10:04  
Job time : 61.2737 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: September 24, 2004, 07:28:32 ; Search time 9.68421 Seconds  
(without alignments)  
116.704 Million cell updates/sec

Title: US-09-498-556C-357  
Perfect score: 11  
Sequence: 1 LXXR 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Jan04.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	81.8	3	AAW84193	AAW84193 Peptide C
2	9	81.8	4	AAW48182	AAW48182 Conantoki
3	9	81.8	4	AAW49864	AAW49864 Conantoki
4	9	81.8	4	AAW79019	AAW79019 Amino aci
5	9	81.8	4	AAW51957	AAW51957 PSA antib
6	9	81.8	4	AAW51953	AAW51953 PSA antib
7	9	81.8	4	AAW24457	AAW24457 BoNT/A N-
8	9	81.8	4	AAW14417	AAW14417 C-termina
9	9	81.8	4	AAW20561	AAW20561 Soybean d
10	9	81.8	4	AAW13761	AAW13761 Novel hum
11	9	81.8	4	AAW36682	AAW36682 G protein
12	9	81.8	4	AAW36806	AAW36806 G protein
13	9	81.8	4	AAW17690	AAW17690 Type IV c
14	9	81.8	4	AAW97797	AAW97797 MAPKK-2 g
15	9	81.8	5	AAW03446	AAW03446 Accessory
16	9	81.8	5	AAW48999	AAW48999 Sequence
17	9	81.8	5	AAW50136	AAW50136 Ovine gro
18	9	81.8	5	AAW50134	AAW50134 Ovine gro
19	9	81.8	5	AAW50135	AAW50135 Ovine gro
20	9	81.8	5	AAW77333	AAW77333 Human apo
21	9	81.8	5	AAW00252	AAW00252 Subtilisi
22	9	81.8	5	AAW00251	AAW00251 Subtilisi
23	9	81.8	5	AAW09696	AAW09696 Cyclic pe
24	9	81.8	5	AAW09695	AAW09695 Cyclic pe
25	9	81.8	5	AAW37317	AAW37317 Mammalian

26	9	81.8	5	AAW51458	AAW51458 AAV VP3 d
27	9	81.8	5	AAE05461	AAE05461 MASP subs
28	9	81.8	5	AAE11109	AAE11109 Tryptic p
29	9	81.8	5	AAE11144	AAE11144 Chymotryp
30	9	81.8	5	AAW51956	AAW51956 PSA antib
31	9	81.8	5	AAW51952	AAW51952 PSA antib
32	9	81.8	5	ABG77841	ABG77841 Targettin
33	9	81.8	5	AAE26308	AAE26308 Human rBP
34	9	81.8	5	ABP55342	ABP55342 Spinacia
35	9	81.8	5	ABP55341	ABP55341 Cucurbita
36	9	81.8	5	ABP75051	ABP75051 Proteome
37	9	81.8	5	ABR55078	ABR55078 MMP subst
38	9	81.8	5	ADC17431	ADC17431 Type IV c
39	9	81.8	6	AAW46808	AAW46808 Phytase d
40	9	81.8	6	AAW50143	AAW50143 Ovine gro
41	9	81.8	6	AAW50142	AAW50142 Ovine gro
42	9	81.8	6	AAW50141	AAW50141 Ovine gro
43	9	81.8	6	AAW55741	AAW55741 dsRNA-dep
44	9	81.8	6	AAW90511	AAW90511 Hydridoma
45	9	81.8	6	AAW74602	AAW74602 3-methylc

ALIGNMENTS

RESULT 1  
AAW84193  
ID AAW84193 standard; peptide; 3 AA.

AC AAW84193;

DT 25-MAR-1999 (first entry)

DE Peptide comprising a proteinase site.

XX Proteinase site; bone morphogenetic fusion protein; bone binding site;

XX bone morphogenetic protein; transforming growth factor beta;

XX active fragment; wound healing; bone growth.

XX Unidentified.

XX W09855137-A1.

PD 10-DEC-1998.

XX 02-JUN-1998; 98WO-US011189.

XX 03-JUN-1997; 97US-00868452.

XX (NIMN/) NIMNI M E.

XX (HALL/) HALL F L.

XX (WULL/) WU L.

XX (SHOR/) SHORS E C.

XX Nimni ME, Hall FL, Wu L, Han B, Shors EC;

XX WPI; 1999-059875/05.

XX New bone morphogenetic fusion proteins - comprising a purification tag and a bone morphogenetic active fragment, used for enhancing wound healing or bone growth.

PS Claim 8; Page 34; 64pp; English.

XX The present peptide represents a proteinase site used in the creation of the bone morphogenetic fusion proteins of the invention. The bone morphogenetic fusion protein may contain some or all of the following elements: a purification tag, a proteinase site, an ECM/bone binding site, a second proteinase site, and a bone morphogenetic protein active fragment. The fusion proteins of the invention also includes proteins that have transforming growth factor beta active fragments instead of bone morphogenetic protein active fragments. The bone morphogenetic

CC fusion proteins can be used for enhancing wound healing or bone growth  
 XX  
 SQ Sequence 3 AA;  
 Query Match 81.8%; Score 9; DB 2; Length 3;  
 Best Local Similarity 66.7%; Pred. No. 1.3e+06;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 LXR 3  
 |  
 |  
 1 LTR 3  
 Db  
 RESULT 2  
 AAW48182  
 ID AAW48182 standard; peptide; 4 AA.  
 XX AC AAW48182;  
 XX DT 30-JUN-1998 (first entry)  
 XX DE Conantokin peptide derivative.  
 XX KW Conantokin; predatory cone snail; treatment; neurologic disorder;  
 KW psychiatric disorder; anticonvulsant; neuroprotective;  
 KW analgesic. HIV infection; ophthalmic indication; memory; learning defect;  
 KW cognitive defect.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 FT Modified-site 4 /note= "gamma-carboxyglutamic acid"  
 FT FT  
 PN WO9803541-A1.  
 XX PD 29-JAN-1998.  
 XX PF 21-JUL-1997; 97WO-US012618.  
 XX PR 22-JUL-1996; 96US-00684742.  
 XX PA (UTAH ) UNIV UTAH RES FOUND.  
 PA (COGN-) COGNETIX INC.  
 XX PI Abogadie FC, Cruz LJ, Olivera BM, Walker C, Colledge C;  
 PI Hillyard DR, Jimenez E, Laver RT, Zhou L, Shen GS, McCabe RT;  
 PI Rivier JE;  
 XX WPI; 1998-120694/11.  
 XX  
 XX New conantokin peptide(s) - useful for e.g. treating neurologic or  
 PT psychiatric disorders, or the management of pain.  
 XX  
 XX Claim 15; Page 98; 122pp; English.  
 XX  
 XX The present sequence is a conantokin peptide derivative, which can be  
 CC used to treat neurologic and psychiatric disorders, e.g. as an  
 CC anticonvulsant, neuroprotective or analgesic agent. Neurologic and  
 CC psychiatric disorders include epilepsy, convulsions, neurotoxic injury  
 CC (associated with conditions of hypoxia, anoxia or ischaemia, which  
 CC typically follow stroke, cerebrovascular accident, brain or spinal cord  
 CC trauma, myocardial infarct, physical trauma, drowning, suffocation,  
 CC perinatal asphyxia or hypoglycaemic events), neurodegeneration  
 CC (associated with Alzheimer's disease, senile dementia, Amyotrophic  
 CC lateral sclerosis, Multiple Sclerosis, Parkinson's disease, Huntington's  
 CC disease, Down's Syndrome, Korsakoff's disease, schizophrenia, AIDS  
 CC dementia, multi-infarct dementia, Binswanger dementia and neuronal damage  
 CC associated with uncontrolled seizures), chemical toxicity (such as  
 CC addition, and morphine, opiate, opioid and barbiturate tolerance), pain  
 CC (acute, chronic, migraine), anxiety, major depression, manic-depressive  
 CC illness, obsessive-compulsive disorder, schizophrenia and mood disorders  
 CC (such as bipolar disorder, unipolar depression, dysthymia and seasonal

CC effective disorder) and dystonia (movement disorder), sleep disorder,  
 CC muscle relaxation and urinary incontinence. The peptide can also be used  
 CC to treat HIV infection, ophthalmic indication and memory, learning or  
 CC cognitive defects  
 XX SQ Sequence 4 AA;  
 Query Match 81.8%; Score 9; DB 2; Length 4;  
 Best Local Similarity 66.7%; Pred. No. 1.3e+06;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 LXR 3  
 |  
 |  
 1 LAR 3  
 Db  
 RESULT 3  
 AAW49964  
 ID AAW49964 standard; peptide; 4 AA.  
 XX AC AAW49964;  
 XX DT 30-JUN-1998 (first entry)  
 XX DE Conantokin peptide derivative.  
 XX KW Conantokin; predatory cone snail; treatment; neurologic disorder;  
 KW psychiatric disorder; anticonvulsant; neuroprotective;  
 KW analgesic. HIV infection; ophthalmic indication; memory; learning defect;  
 KW cognitive defect.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 FT Modified-site 4 /note= "gamma-carboxyglutamic acid"  
 FT FT  
 PN WO9803189-A1.  
 XX PD 29-JAN-1998.  
 XX PF 21-JUL-1997; 97WO-US012652.  
 XX PR 22-JUL-1996; 96US-00684750.  
 PR 06-DEC-1996; 96US-00762377.  
 XX PA (COGN-) COGNETIX INC.  
 XX PI McCabe RT, Zhou L, Laver RT;  
 XX WPI; 1998-120469/11.  
 XX  
 XX Use of conantokin peptide(s) - for treating disorders involving excessive  
 PT excitation of nerve cells by excitatory amino acids or agonists of the N-  
 PT methyl-D-aspartate receptor.  
 XX  
 XX Claim 27; Page 69; 122pp; English.  
 XX  
 XX The present sequence is a conantokin peptide derivative, which can be  
 CC used to treat neurologic and psychiatric disorders, e.g. as an  
 CC anticonvulsant, neuroprotective or analgesic agent. Neurologic and  
 CC psychiatric disorders include epilepsy, convulsions, neurotoxic injury  
 CC (associated with conditions of hypoxia, anoxia or ischaemia, which  
 CC typically follow stroke, cerebrovascular accident, brain or spinal cord  
 CC trauma, myocardial infarct, physical trauma, drowning, suffocation,  
 CC perinatal asphyxia or hypoglycaemic events), neurodegeneration  
 CC (associated with Alzheimer's disease, senile dementia, Amyotrophic  
 CC lateral sclerosis, Multiple Sclerosis, Parkinson's disease, Huntington's  
 CC disease, Down's Syndrome, Korsakoff's disease, schizophrenia, AIDS  
 CC dementia, multi-infarct dementia, Binswanger dementia and neuronal damage  
 CC associated with uncontrolled seizures), chemical toxicity (such as  
 CC addition, and morphine, opiate, opioid and barbiturate tolerance), pain  
 CC (acute, chronic, migraine), anxiety, major depression, manic-depressive  
 CC illness, obsessive-compulsive disorder, schizophrenia and mood disorders  
 CC (such as bipolar disorder, unipolar depression, dysthymia and seasonal

CC illness, obsessive-compulsive disorder, schizophrenia and mood disorders  
CC (such as bipolar disorder, unipolar depression, dysthymia and seasonal  
CC effective disorder) and dystonia (movement disorder) sleep disorder.  
CC muscle relaxation and urinary incontinence. The peptide can also be used  
CC to treat HIV infection, ophthalmic indication and memory, learning or  
CC cognitive defects  
XX  
SQ Sequence 4 AA;

Query Match 81.8%; Score 9; DB 2; Length 4;  
Best Local Similarity 66.7%; Pred. No. 1.3e+06;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3  
|  
|  
Db 1 LAR 3

RESULT 4  
AAG79019  
ID AAG79019 standard; peptide; 4 AA.

XX AAG79019;

DT 10-DEC-2001 (first entry)

DE Amino acid sequence of conantokin R domain III.

XX Conantokin; cone snail; nerve cell excitation; NMDA receptor; epileps;  
KW N-methyl-D-aspartate receptor; pain; psychiatric disorder;  
KW neurotoxic injury; hypoxia; anoxia; ischemia; neurodegeneration;  
KW chemical toxicity; addiction; drug craving; psychiatric disorder;  
KW anxiety; depression; obsessive compulsive disorder; schizophrenia;  
KW mood disorder; ophthalmic disorder; neurological disorder; dystonia;  
KW sleep disorder; muscle relaxation; urinary incontinence;  
KW cognition enhancement; HIV infection.

XX Conus radiatus.

XX Key Location/Qualifiers

FT Modified-site 4 /note= "gamma-carboxyglutamic acid"

XX US6277825-B1.

XX 21-AUG-2001.

XX 20-JUL-1999; 98US-00357141.

XX 22-JUL-1996; 98US-00684750.

XX 06-DEC-1996; 98US-00762377.

XX 21-JUL-1997; 97MO-US012652.

XX 10-FEB-1999; 98US-00142076.

XX 01-APR-1999; 99US-00283277.

XX (UTAH ) UNIV UTAH RES FOUND.

XX (COGN-) COGNEX INC.

XX Olivera BM, McIntosh JM, McCabe RT, Layer RT, Zhou L;

XX WPI; 2001-601377/68.

XX Use of conantokin peptide or its derivatives or a conantokin peptide

XX chimera for treating disorders e.g. migraine.

XX Claim 9; Col 80; 60pp; English.

XX AAG79012-43 and AAG790054-56 represent domains of conantokin peptides.

XX Conantokin differ from conotoxins, in that they contain gamma-

XX carboxyglutamic acid. The conantokin are derived from the venom of cone

XX snails. They are used for the treatment of disorders in which the

XX pathophysiology involves excessive excitation of nerve cells by

XX excitatory amino acids or agonist of N-methyl-D-aspartate (NMDA)

CC receptor. The conantokin peptides are used for the treatment of disorders  
CC such as pain; neurologic or psychiatric disorders such epileps; for  
CC reducing neurotoxic injury associated with conditions of hypoxia, anoxia  
CC or ischemia; for treating neurodegeneration; for treating chemical,  
CC toxicity such as addiction, drug craving, alcohol abuse, morphine, opioid  
CC and barbiturate tolerance; for treating psychiatric disorders such as  
CC anxiety, major depression, manic-depression illness, obsessive compulsive  
CC disorder, schizophrenia or mood disorder; for treating ophthalmic  
CC disorder; for treating additional neurological disorders e.g. dystonia,  
CC sleep disorder, muscle relaxation and urinary incontinence; for  
CC memory/cognition enhancement; for treating HIV infection

XX Sequence 4 AA;

Query Match 81.8%; Score 9; DB 4; Length 4;  
Best Local Similarity 66.7%; Pred. No. 1.3e+06;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3  
|  
|  
Db 1 LAR 3

RESULT 5

AAMS1957  
ID AAMS1957 standard; peptide; 4 AA.

XX AAMS1957;

XX 01-FEB-2002 (first entry)

DE PSA antibody preparation immunogen peptide #9.

XX PSA; prostate-specific antigen; prostatic carcinoma; pro-kallikrein 2;  
KW human; antibody; immunogen.

XX Homo sapiens.

XX DE10032040-A1.

XX 29-NOV-2001.

XX 05-JUL-2000; 2000DE-01032040.

XX 24-MAY-2000; 2000DE-01025387.

XX (HOFF ) ROCHE DIAGNOSTICS GMBH.

XX WPI; 2002-042633/06.

XX Antibody comprising specificity for some truncated forms of pro-prostate  
PT specific antigen, useful in immunoassays for differential diagnosis of  
PT prostatic carcinoma.

XX Disclosure; Page 11; 16pp; German.

XX The present invention relates to an antibody specific for (-5,-6 and -7)-  
CC PROPSA (prostate-specific antigen), but with no reactivity with (-4) or  
CC shorter forms of the protein. The antibody can be used to diagnose  
CC prostatic carcinoma and to detect human pro-kallikrein 2. The present  
CC sequence is a peptide which forms part of a peptide-containing immunogen  
CC described in the exemplification of the invention

XX Sequence 4 AA;

Query Match 81.8%; Score 9; DB 5; Length 4;  
Best Local Similarity 66.7%; Pred. No. 1.3e+06;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3  
|  
|  
Db 1 LSR 3

RESULT 6  
AAM51953  
ID AAM51953 standard; peptide; 4 AA.  
XX AC AAM51953;  
XX DT 01-FEB-2002 (first entry)  
XX DE PSA antibody preparation immunogen peptide #5.  
XX KW PSA; prostate-specific antigen; prostatic carcinoma; pro-kallikrein 2;  
XX KW human; antibody; immunogen.  
XX OS Homo sapiens.  
XX PN DE10032040-A1.  
XX PD 29-NOV-2001.  
XX PF 05-JUL-2000; 2000DE-01032040.  
XX PR 24-MAY-2000; 2000DE-01025387.  
XX PA (HOFF ) ROCHE DIAGNOSTICS GMBH.  
XX DR WPI; 2002-042633/06.  
XX PT Antibody comprising specificity for some truncated forms of pro-prostate  
PT specific antigen, useful in immunoassays for differential diagnosis of  
PT prostatic carcinoma.  
XX PS Disclosure; Page 11; 16pp; German.  
XX CC The present invention relates to an antibody specific for (-5,-6 and -7) -  
CC proPSA (prostate-specific antigen), but with no reactivity with (-4) or  
CC shorter forms of the protein. The antibody can be used to diagnose  
CC prostatic carcinoma and to detect human pro-kallikrein 2. The present  
CC sequence is a peptide which forms part of a peptide-containing immunogen  
CC described in the exemplification of the invention  
XX SQ Sequence 4 AA;  
Query Match 81.8%; Score 9; DB 5; Length 4;  
Best Local Similarity 66.7%; Pred. No. 1.3e+06;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 LXR 3  
Db 2 LSR 4  
RESULT 7  
AAE24457  
ID AAE24457 standard; peptide; 4 AA.  
XX AC AAE24457;  
XX DT 04-OCT-2002 (first entry)  
XX DE BONT/A N-glycosylation site #1.  
XX KW Neurotoxin; biological persistence; blepharospasm; pain; therapy;  
KW neuromuscular disorder; cervical dystonia; oromandibular dystonia;  
KW spasmodic dysphonia; laryngeal dystonia; muscular tension; asthma;  
KW neuralgia; autonomic nervous system disorder; sweating; salivation;  
KW headache; neuropathy; botulinum toxin serotype A; BONT/A.  
XX OS Clostridium botulinum.  
XX PN WO200240506-A2.  
XX PD 23-MAY-2002.

XX PF 16-NOV-2001; 2001WO-US044030.  
XX PR 17-NOV-2000; 2000US-0249540P.  
XX PA (ALLR ) ALLERGAN SALES INC.  
XX PI Steward LE, Spanoyannis A, Lin W, Aoki KR;  
XX DR WPI; 2002-479904/51.  
XX PT Modified neurotoxin especially Clostridial toxins, useful for treating  
PT neuromuscular and autonomic nervous system disorder and pain, comprises  
PT structural modification to alter biological persistence of neurotoxin.  
XX PS Disclosure; Page 19; 55pp; English.  
XX CC The invention related to modified neurotoxins especially Clostridial  
CC botulinum toxins with altered biological persistence. These toxins  
CC comprise a structural modification which is effective to alter the  
CC biological persistence. Modified neurotoxins of the invention are used  
CC for treating biological disorders which include neuromuscular disorder  
CC e.g. strabismus, blepharospasm, spasmodic torticollis (cervical  
CC dystonia), oromandibular dystonia and spasmodic dysphonia (laryngeal  
CC dystonia), autonomic nervous system disorders e.g. excessive salivation  
CC and sweating, asthma etc. and pain e.g. headache, muscular tension,  
CC neuralgia and neuropathy. The present sequence is botulinum toxin  
CC serotype A (BONT/A) N-glycosylation site. This peptide is derived from C.  
CC botulinum  
XX SQ Sequence 4 AA;  
Query Match 81.8%; Score 9; DB 5; Length 4;  
Best Local Similarity 66.7%; Pred. No. 1.3e+06;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 LXR 3  
Db 2 LTR 4  
RESULT 8  
AAE14417  
ID AAE14417 standard; peptide; 4 AA.  
XX AC AAE14417;  
XX DT 26-MAR-2002 (first entry)  
XX DE C-terminal sequence of an arginine-containing peptide.  
XX KW Protein quantification; protein identification; signature peptide;  
KW mass spectrometric analysis; proteolytic peptide; affinity ligand;  
KW arginine; isobaric peptide.  
XX OS Unidentified.  
XX PN WO200186306-A2.  
XX PD 15-NOV-2001.  
XX PF 04-MAY-2001; 2001WO-US014418.  
XX PR 05-MAY-2000; 2000US-0203227P.  
XX PR 31-MAY-2000; 2000US-0208184P.  
XX PR 31-MAY-2000; 2000US-0208372P.  
XX PA (PURD ) PURDUE RES FOUND.  
PA (REGN ) REGNIER P E.  
PA (CHAK ) CHAKRABORTY A B.  
PA (DORM ) DORMADY S J.  
PA (GENG ) G'ENG M.  
PA (JICJ ) JI J.

PA (RIGG/) RIGGS L D.  
 PA (SIOM/) SIOMA C S.  
 PA (WANG/) WANG S.  
 PA (ZHAN/) ZHANG X.  
 XX Regnier FE, Chakraborty AB, Dormady SJ, G'eng M, Ji J, Riggs LD;  
 PI Sioma CS, Wang S, Zhang X;  
 XX WPI; 2002-089810/12.  
 XX  
 XX Analyzing differences in protein content in protein samples, useful for  
 PT identifying protein(s) in a sample or multiple proteins in a single  
 PT complex mixture, by employing mass spectrometric analysis of signature  
 PT peptides.  
 XX  
 XX Example 9; Page 73; 106pp; English.  
 PS  
 XX The invention relates to analysing differences in protein content among  
 CC plural protein samples comprising employing mass spectrometric analysis  
 CC of proteolytic peptide fragments. The method comprises fragmenting at  
 CC least a first protein sample and a second protein sample to produce a  
 CC first peptide pool and a second peptide pool; isotopically labelling at  
 CC least a portion of the peptides in at least one of the pools so as to  
 CC permit resolution of otherwise identical peptides in the first and second  
 CC peptide pools by mass analysis; contacting peptides from at least a  
 CC portion of both of the peptide pools with a capture moiety to yield  
 CC affinity-selected peptides comprising an affinity ligand; where the  
 CC capture moiety selects for the affinity ligand; and analysing the  
 CC affinity-selected peptides by mass spectrometry to determine one or more  
 CC differences between the first and second samples. The method of the  
 CC invention is useful for identifying one or more proteins in a sample,  
 CC and is typically useful for identifying multiple proteins in a single  
 CC complex mixture. The method is also useful for quantifying proteins in a  
 CC sample or several samples. The advantages of the novel method are that it  
 CC is easier to separate peptides than proteins; native structure of the  
 CC protein does not have to be maintained during the analysis; structural  
 CC variants do not interfere; and putative proteins suggested from DNA  
 CC database can be recognised by using a signature peptide probe. The  
 CC present method permits identification of a protein in a complex sample  
 CC without purifying the protein or obtaining its composite peptide  
 CC signature. The present sequence is C-terminal fragment of a  
 CC differentially acetylated arginine-containing peptide which is used to  
 CC generate fragment ions for distinguishing isobaric peptides  
 XX  
 SQ Sequence 4 AA;  
 Query Match 81.8%; Score 9; DB 5; Length 4;  
 Best Local Similarity 66.7%; Pred. No. 1.3e+06;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 LXR 3  
 |  
 2 LAR 4  
 DB  
 RESULT 9  
 AAE20561  
 ID AAE20561 standard; peptide; 4 AA.  
 XX  
 AC AAE20561;  
 XX  
 XX 01-JUL-2002 (first entry)  
 DT  
 XX Soybean diverged delta-9 fatty acid desaturase peptide #7.  
 DE  
 XX Soybean; diverged delta-9 fatty acid desaturase; enzyme; cellular lipid;  
 KW oil; transgenic plant; gene mapping; immunisation.  
 KW  
 XX Glycine max.  
 OS  
 XX WC020216565-A2.  
 PN  
 XX 28-FEB-2002.  
 PD

XX 22-AUG-2001; 2001WO-US026246.  
 PF  
 XX 22-AUG-2000; 2000US-0226996P.  
 PR  
 XX (DUPO ) DU PONT DE NEMOURS & CO E I.  
 PA  
 XX Booth JR, Cahoon RE, Hitz WD, Kinney AJ, Yadav NS;  
 PI  
 XX WPI; 2002-269353/31.  
 DR  
 XX New delta-9 fatty acid desaturase polypeptides and polynucleotides,  
 PT useful in creating transgenic plants having altered levels of mono-, poly  
 PT - and unsaturated fatty acids and in increasing the unsaturation levels  
 PT in cellular lipids.  
 XX  
 XX Claim 12; Page 48; 77pp; English.  
 PS  
 XX The present invention relates to diverged delta-9 fatty acid desaturase  
 CC proteins and polynucleotides encoding such proteins. The nucleic acid  
 CC sequences may be used to increase the level of unsaturation in cellular  
 CC lipids, including oil, in tissues when the enzyme is absent or rate-  
 CC limiting, to isolate cDNAs and genes encoding homologous proteins from  
 CC the same or other plant species and to create transgenic plants in which  
 CC the polypeptides are present at higher or lower levels than normal or in  
 CC cell types or developmental stages in which they are not normally found,  
 CC thus altering the level of mono-, poly- and unsaturated fatty acids in  
 CC those cells. They are useful as probes for genetic and physical gene  
 CC mapping and as markers, e.g. restriction fragment length polymorphism  
 CC (RFLP) markers. The peptides can be used to immunise animals to produce  
 CC antibodies specific for the peptides and proteins. The present sequence  
 CC is soybean diverged delta-9 fatty acid desaturase peptide (residues 256-  
 CC 259)  
 XX  
 SQ Sequence 4 AA;  
 Query Match 81.8%; Score 9; DB 5; Length 4;  
 Best Local Similarity 66.7%; Pred. No. 1.3e+06;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 LXR 3  
 |  
 2 LAR 4  
 DB  
 RESULT 10  
 ABU13761  
 ID ABU13761 standard; peptide; 4 AA.  
 XX  
 AC ABU13761;  
 XX  
 XX 25-FEB-2003 (first entry)  
 DT  
 XX Novel human transporter protein related peptide #4.  
 DE  
 XX Human; gamma-aminobutyric acid; GABA; neurotransmitter transporter;  
 KW transgenic animal; ribozyme design; drug screening; gene therapy.  
 KW  
 XX Homo sapiens.  
 OS  
 XX US2002142381-A1.  
 PN  
 XX 03-OCT-2002.  
 PD  
 XX 28-MAR-2001; 2001US-00818656.  
 PF  
 XX 28-MAR-2001; 2001US-00818656.  
 PR  
 XX (GONG/) GONG F.  
 PA (KETC/) KETCHUM K A.  
 PA (DFRA/) DI FRANCESCO V.  
 PA (BEAS/) BEASLEY E M.  
 XX

PI Gong F, Ketchum KA, Di Francesco V, Beasley EM;  
 XX WPI; 2003-102517/09.  
 XX  
 XX Novel human transporter protein, related to gamma-aminobutyric acid  
 PT neurotransmitter transporter subfamily useful as model for developing  
 PT human therapeutic targets and serves as target for human therapeutics.  
 XX  
 XX Disclosure; Page 45; 114pp; English.  
 XX  
 XX The invention describes an isolated human transporter peptide (I) that is  
 CC related to the gamma-aminobutyric acid (GABA) neurotransmitter  
 CC transporter subfamily. (I) and the nucleic acid encoding it (II) can be  
 CC used as models for the development of human therapeutic targets, aid in  
 CC the identification of therapeutic proteins and serve as targets for the  
 CC development of human therapeutic agents that modulate transporter  
 CC activity. (I) is used to raise antibodies or to elicit another immune  
 CC response, as a reagent in assays designed to quantitatively determine  
 CC levels of the protein in biological fluids, and as markers for tissues in  
 CC which the corresponding protein is preferentially expressed. The  
 CC transporter proteins isolated from humans and their human/mammalian  
 CC orthologues serve as targets for identifying agents for use in mammalian  
 CC therapeutic applications, and biological assays related to transporter  
 CC proteins that are related to members of the GABA neurotransmitter  
 CC transporter subfamily. The proteins and peptides also provide a target  
 CC for diagnosing a disease or predisposition to disease mediated by the  
 CC peptide, and are useful for treating a disorder characterised by absence  
 CC of inappropriate unwanted or altered expression of the protein. The  
 CC antibodies are also useful for assessing normal and aberrant subcellular  
 CC localisation of cells in various tissues in an organism, in  
 CC pharmacogenomic analysis, for tissue typing and for inhibiting protein  
 CC function. (III) is useful for constructing recombinant vectors, host cells  
 CC and transgenic animals; for designing ribozymes; in drug screening; in  
 CC diagnostic assays for qualitative changes in gene expression,  
 CC particularly in qualitative changes that lead to pathology; in gene  
 CC therapy; and to detect mutations in genes encoding transporters. This is  
 CC the amino acid sequence of a polypeptide related to the novel human  
 CC aminobutyric acid (GABA) transporter related protein  
 XX  
 XX Sequence 4 AA;  
 SQ  
 Query Match 81.8%; Score 9; DB 6; Length 4;  
 Best Local Similarity 66.7%; Pred. No. 1.3e+06;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 LXR 3  
 Db 2 LTR 4  
 RESULT 11  
 ABJ36682  
 ID ABJ36682 standard; peptide; 4 AA.  
 XX  
 XX ABJ36682;  
 XX  
 XX 01-MAY-2003 (first entry)  
 DT  
 XX  
 XX G protein coupled receptor related peptide SEQ ID No 5.  
 DE  
 XX  
 XX Nootropic; cardiant; antiarteriosclerotic; hypotensive; cytostatic;  
 KW antibacterial; analgesic; antiallergic; antiasthmatic; antiinflammatory;  
 KW osteopathic; neuroprotective; anxiolytic; anorectic; lead compound;  
 KW G protein coupled receptor signaling inhibitor; GPCR; library;  
 KW high throughput screening assay; stroke; myocardial infarction;  
 KW restenosis; atherosclerosis; hypotension; cancer; infection; asthma;  
 KW septic shock; pain; allergic disorder; inflammatory bowel disease;  
 KW osteoporosis; obesity; psychotic; neurological disorder; anxiety;  
 KW schizophrenia; Alzheimer's disease.  
 XX Mammalia sp.  
 OS  
 XX Mammalia sp.  
 PN W0200272778-A2.

XX 19-SEP-2002.  
 PD  
 XX 14-MAR-2002; 2002WO-US007561.  
 XX  
 XX 14-MAR-2001; 2001US-0275472P.  
 PR  
 XX 11-MAY-2001; 2001US-00852910.  
 PR  
 XX (CUEB-) CUE BIOTECH.  
 FA  
 XX Gilchrist A, Hamm HE;  
 FI  
 XX WPI; 2003-247841/24.  
 DR  
 XX Identifying G protein coupled receptor (GPCR) signaling inhibitors,  
 PT useful in screening drugs for treating stroke, cancers or pain, by  
 PT identifying compounds that block GPCR mediated signaling with high  
 PT affinity and specificity.  
 XX  
 XX Disclosure; Page 12; 94pp; English.  
 PS  
 XX The invention relates to a novel method for identifying a G protein  
 CC coupled receptor (GPCR) signaling inhibitor. The novel method comprises  
 CC selecting or identifying a member of a library of peptides and/or  
 CC candidate compounds, having binding to a GPCR of higher affinity than  
 CC that of the native peptide. The peptide library is based on a native GPCR  
 CC binding peptide. The method is useful for identifying inhibitors of a G  
 CC protein coupled receptor (GPCR) signaling. The method is particularly  
 CC useful for identifying drugs that antagonise the binding between a GPCR  
 CC and its extracellular ligand(s). The method is especially useful in  
 CC modern high throughput screening assays for identifying potent lead  
 CC compounds. The compounds, peptides or inhibitors identified by the method  
 CC are useful for preventing, ameliorating or treating diseases in which  
 CC GPCR signaling is a causative factor or in which a specific class of G  
 CC protein is relevant, e.g. stroke, myocardial infarction, restenosis,  
 CC atherosclerosis, hypotension, cancers, infections, septic shock, pain,  
 CC allergic disorders, asthma, inflammatory bowel disease, osteoporosis,  
 CC obesity, or psychotic and neurological disorders (e.g. anxiety,  
 CC schizophrenia or Alzheimer's disease). This sequence represents a peptide  
 CC relating to the G protein coupled receptors of the invention  
 XX  
 XX Sequence 4 AA;  
 SQ  
 Query Match 81.8%; Score 9; DB 6; Length 4;  
 Best Local Similarity 66.7%; Pred. No. 1.3e+06;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 LXR 3  
 Db 1 LSR 3  
 RESULT 12  
 ABJ36806  
 ID ABJ36806 standard; peptide; 4 AA.  
 XX  
 XX ABJ36806;  
 AC  
 XX 01-MAY-2003 (first entry)  
 DT  
 XX  
 XX G protein coupled receptor related peptide SEQ ID No 155.  
 DE  
 XX Nootropic; cardiant; antiarteriosclerotic; hypotensive; cytostatic;  
 KW antibacterial; analgesic; antiallergic; antiasthmatic; antiinflammatory;  
 KW osteopathic; neuroprotective; anxiolytic; anorectic; lead compound;  
 KW G protein coupled receptor signaling inhibitor; GPCR; library;  
 KW high throughput screening assay; stroke; myocardial infarction;  
 KW restenosis; atherosclerosis; hypotension; cancer; infection; asthma;  
 KW septic shock; pain; allergic disorder; inflammatory bowel disease;  
 KW osteoporosis; obesity; psychotic; neurological disorder; anxiety;  
 KW schizophrenia; Alzheimer's disease.  
 XX Unidentified.  
 OS

XX WO200272778-A2.  
 XX 19-SEP-2002.  
 XX 14-MAR-2002; 2002WO-US007561.  
 XX 14-MAR-2001; 2001US-0275472P.  
 XX 11-MAY-2001; 2001US-00852910.  
 XX (CUEB-) CUE BIOTECH.  
 XX Gilchrist A, Hamm HE;  
 XX WPI; 2003-247841/24.  
 XX Identifying G protein coupled receptor (GPCR) signaling inhibitors,  
 XX useful in screening drugs for treating stroke, cancers or pain, by  
 XX identifying compounds that block GPCR mediated signaling with high  
 XX affinity and specificity.  
 XX Example 8; Page 178; 94pp; English.  
 XX The invention relates to a novel method for identifying a G protein  
 XX coupled receptor (GPCR) signaling inhibitor. The novel method comprises  
 XX selecting or identifying a member of a library of peptides and/or  
 XX candidate compounds, having binding to a GPCR of higher affinity than  
 XX that of the native peptide. The peptide library is based on a native GPCR  
 XX binding peptide. The method is useful for identifying inhibitors of a G  
 XX protein coupled receptor (GPCR) signaling. The method is particularly  
 XX useful for identifying drugs that antagonise the binding between a GPCR  
 XX and its extracellular ligand(s). The method is especially useful in  
 XX modern high throughput screening assays for identifying potent lead  
 XX compounds. The compounds, peptides or inhibitors identified by the method  
 XX are useful for preventing, ameliorating or treating diseases in which  
 XX GPCR signaling is a causative factor or in which a specific class of G  
 XX protein is relevant, e.g. stroke, myocardial infarction, restenosis,  
 XX atherosclerosis, hypotension, cancers, infections, septic shock, pain,  
 XX allergic disorders, asthma, inflammatory bowel disease, osteoporosis,  
 XX obesity, or psychotic and neurological disorders (e.g. anxiety,  
 XX schizophrenia or Alzheimer's disease). This sequence represents a peptide  
 XX relating to the G protein coupled receptors of the invention  
 XX  
 XX Sequence 4 AA;  
 XX  
 XX Query Match 81.8%; Score 9; DB 6; Length 4;  
 XX Best Local Similarity 66.7%; Pred. NO. 1.3e+06;  
 XX Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 XX  
 XX QY 1 LXR 3  
 XX DB 2 LTR 4  
 XX  
 XX RESULT 13  
 XX ADC17690  
 XX ID ADC17690 standard; peptide; 4 AA.  
 XX AC ADC17690;  
 XX  
 XX 18-DEC-2003 (first entry)  
 XX  
 XX Type IV collagen NC1 domain related peptide SEQ ID NO:295.  
 XX  
 XX crystallised NC1 domain hexamer of type IV collagen;  
 XX angiogenesis inhibitor; angiogenesis-mediated disease;  
 XX tumour metastasis inhibitor; tumour growth inhibitor;  
 XX endothelial cell interaction inhibitor;  
 XX basal lamina membrane formation inhibitor; cytostatic; antiapoptotic;  
 XX antianemic; ophthalmological; antiarteriosclerotic; antiulcer;  
 XX endothelial cell adhesion inhibitor;  
 XX endothelial cell proliferation inhibitor; glaucoma; sickle cell anaemia;  
 XX ulcerative colitis; psoriasis; atherosclerosis; rheumatoid arthritis;

KW blood-borne tumour.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PN WO2003012122-A2.  
 XX  
 XX 13-FEB-2003.  
 XX  
 XX 26-JUL-2002; 2002WO-US023763.  
 XX  
 XX 27-JUL-2001; 2001US-0308523P.  
 XX 29-OCT-2001; 2001US-0351289P.  
 XX 22-MAR-2002; 2002US-0366854P.  
 XX 03-JUN-2002; 2002US-0385362P.  
 XX  
 XX (UNIV ) UNIV KANSAS MEDICAL CENT.  
 XX (SUND/) SUNDARAMOORTHY M.  
 XX (HUDS/) HUDSON B.  
 XX Sundaramoorthy M, Hudson B;  
 XX WPI; 2003-332730/31.  
 XX  
 XX New polypeptide, useful for treating an angiogenesis-mediated disease or  
 XX condition consisting of glaucoma or blood-borne tumors or for inhibiting  
 XX basal lamina membrane formation in cell or tissue development.  
 XX  
 XX Claim 57; SEQ ID NO 295; 160pp; English.  
 XX  
 XX The present invention describes a crystallised NC1 domain hexamer of type  
 XX IV collagen (I). Also described: (1) a chimeric polypeptide; (2) a  
 XX pharmaceutical composition comprising the polypeptide and a carrier; (3)  
 XX inhibiting angiogenesis in tissue; (4) treating an angiogenesis-mediated  
 XX disease or condition in a mammal; (5) inhibiting tumour metastasis or  
 XX growth; (5) inhibiting endothelial cell interaction with the  
 XX extracellular matrix in an animal tissue; (6) inhibiting basal lamina  
 XX membrane formation in cell or tissue development; (7) a crystal of an NC1  
 XX domain hexamer of type IV collagen; (8) identifying inhibitors of type IV  
 XX collagen assembly; and (9) an inhibitor of type IV collagen assembly. A  
 XX crystallised NC1 domain hexamer of type IV collagen (I) has cytostatic,  
 XX antiproliferative, antianemic, ophthalmological, antiarteriosclerotic and  
 XX anticancer activities, and can be used as an inhibitor of angiogenesis,  
 XX tumour growth, tumour metastasis, endothelial cell adhesion, endothelial  
 XX cell proliferation, and basal lamina assembly. A (I) polypeptide can be  
 XX used for treating an angiogenesis-mediated disease or condition  
 XX consisting of glaucoma, sickle cell anaemia, ulcerative colitis,  
 XX psoriasis, atherosclerosis, rheumatoid arthritis or blood-borne tumours  
 XX or for inhibiting basal lamina membrane formation in cell or tissue  
 XX development. The methods are useful for inhibiting angiogenesis in  
 XX tissue, inhibiting tumour metastasis or growth, inhibiting endothelial  
 XX cell interaction with the extracellular matrix in an animal tissue, and  
 XX identifying inhibitors of type IV collagen assembly. The present sequence  
 XX represents a peptide which is used in the exemplification of the present  
 XX invention.  
 XX  
 XX Sequence 4 AA;  
 XX  
 XX Query Match 81.8%; Score 9; DB 7; Length 4;  
 XX Best Local Similarity 66.7%; Pred. NO. 1.3e+06;  
 XX Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 XX  
 XX QY 1 LXR 3  
 XX DB 1 LAR 3  
 XX  
 XX RESULT 14  
 XX ADC97797  
 XX ID ADC97797 standard; peptide; 4 AA.  
 XX AC ADC97797;  
 XX

DT 01-JAN-2004 (first entry)  
 DE MAPKK-2 gene regulatory peptide.  
 XX gene expression modulation; signalling molecule;  
 XX NF-kappaB/Rel protein inhibitor; antiinflammatory; antiarthritic;  
 KW cerebroprotective; cardiant; antibacterial; immunosuppressive;  
 KW dermatological; nephrotropic; NF-kappaB/Rel protein modulator;  
 KW inflammatory disease; arthritis; ischaemia; cerebrovascular disease;  
 KW ischaemic heart failure; anthrax; angiogenesis; autoimmune disease;  
 KW systemic lupus erythematosus; ulcerative colitis; Addison's disease;  
 KW Goodpasture's disease.  
 XX Synthetic.  
 OS  
 XX WO2003029292-A2.  
 PN  
 XX 10-APR-2003.  
 PD  
 XX 04-OCT-2002; 2002WO-NL0006639.  
 XX 04-OCT-2001; 2001EP-00203748.  
 PR 21-DEC-2001; 2001US-00028075.  
 XX (UYRO-) UNIV ROTTERDAM ERASMUS.  
 PA Khan NA, Benner R;  
 XX WPI; 2003-393380/37.  
 XX  
 XX Modulating gene expression in a cell, useful for treating acute or  
 PT chronic inflammatory diseases (e.g. arthritis), ischemic events or  
 PT autoimmune diseases, comprises providing the cell with a signaling  
 PT molecule.  
 XX  
 PS Disclosure; Page 35; 217pp; English.  
 XX The present invention describes a method for modulating gene expression  
 CC in a cell comprising providing the cell with a signalling molecule  
 CC comprising a peptide or its functional analogue. Also described: (1)  
 CC identifying or obtaining a signalling molecule comprising a peptide or  
 CC its functional derivative or analogue capable of modulating expression of  
 CC a gene in a cell; (2) a signalling molecule useful in modulating  
 CC expression of a gene in a cell and identifiable or obtainable by the  
 CC method of (1); and (3) an inhibitor of nuclear factor (NF)-kappaB/Rel  
 CC protein activation comprising a signalling molecule of (2). The  
 CC signalling molecule has antiinflammatory, antiarthritic,  
 CC cerebroprotective, cardiant, antibacterial, immunosuppressive,  
 CC dermatological and nephrotropic activities, and can be used as a NF-  
 CC kappaB/Rel protein modulator. The signalling molecule is useful for the  
 CC production of a pharmaceutical composition for the modulation of gene  
 CC expression by inhibiting NF-kappaB/Rel protein activation. The method is  
 CC useful for modulating gene expression, and for treating acute or chronic  
 CC inflammatory diseases (e.g. arthritis), ischaemic event including  
 CC cerebrovascular disease and ischaemic heart failure, anthrax,  
 CC angiogenesis, or autoimmune diseases (e.g. systemic lupus erythematosus,  
 CC ulcerative colitis, Addison's disease or Goodpasture's disease). The  
 CC present sequence represents a peptide used in the exemplification of the  
 CC present invention.  
 XX  
 SQ Sequence 4 AA;

Query Match 81.8%; Score 9; DB 7; Length 4;  
 Best Local Similarity 66.7%; Pred. No. 1.3e+06;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3  
 Db 2 LAR 4

RESULT 15  
 AAR03446

ID AAR03446 standard; protein; 5 AA.  
 XX  
 AC AAR03446;  
 XX  
 DT 02-AUG-1990 (first entry)  
 DE  
 XX Accessory moiety derived from transferrin receptor.  
 KW Accessory moiety; transferrin receptor; tumour therapy.  
 XX  
 OS Synthetic.  
 XX EP359347-A.  
 PN  
 XX 21-MAR-1990.  
 PD  
 XX 14-AUG-1989; 89EP-00250014.  
 PF  
 XX 15-AUG-1988; 88US-00232337.  
 PR  
 XX (NEOR-) NEORX CORP.  
 PA Anderson DC, Morgan AC, Abrams PG, Nichols EJ, Fritzberg AR;  
 XX WPI; 1990-085154/12.  
 DR  
 XX Covalently-linked complex for tumour treatment - comprises treating  
 PT protein, cytotoxic agent and enhancing moiety.  
 PT  
 XX Claim 13; Page 22; 23pp; English.  
 PS  
 XX The sequence is one of several possible accessory moieties which can be  
 CC attached to the C-terminus of an anchoring moiety which in turn is linked  
 CC to a targeting protein and a cytotoxic agent. When the anchoring peptide  
 CC intercalates into the target cell plasma membrane the accessory moiety is  
 CC translocated across the membrane and protrudes into the cytoplasm. It can  
 CC then be phosphorylated by cellular kinases at neutral pH. This  
 CC irreversibly anchors the complex to the membrane and may enhance the  
 CC translocation into the cytoplasm. The complex is useful for treatment and  
 CC diagnosis of tumours. See also AAR03435-60  
 XX  
 SQ Sequence 5 AA;

Query Match 81.8%; Score 9; DB 2; Length 5;  
 Best Local Similarity 66.7%; Pred. No. 1.3e+06;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3  
 Db 3 LAR 5

Search completed: September 24, 2004, 07:34:20  
 Job time : 13.6842 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 07:30:36 ; Search time 2.48421 Seconds  
(without alignments)  
83.127 Million cell updates/sec

Title: US-09-498-556C-357

Perfect score: 11

Sequence: 1 LXR4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	81.8	3	1 US-08-470-837-7	Sequence 7, Appli
2	9	81.8	3	4 US-08-868-452-7	Sequence 7, Appli
3	9	81.8	4	1 US-08-288-405A-4	Sequence 4, Appli
4	9	81.8	4	1 US-08-336-343A-36	Sequence 36, Appli
5	9	81.8	4	3 US-09-142-078-19	Sequence 19, Appli
6	9	81.8	4	3 US-09-357-141-19	Sequence 19, Appli
7	9	81.8	4	4 US-09-533-889-19	Sequence 19, Appli
8	9	81.8	4	4 US-09-142-080-19	Sequence 19, Appli
9	9	81.8	4	4 US-09-486-283C-4	Sequence 4, Appli
10	9	81.8	4	4 US-09-535-852-355	Sequence 355, App
11	9	81.8	5	1 US-07-628-923A-12	Sequence 12, Appli
12	9	81.8	5	1 US-08-357-264-5	Sequence 5, Appli
13	9	81.8	5	1 US-08-456-840-38	Sequence 38, Appli
14	9	81.8	5	1 US-08-672-514-5	Sequence 5, Appli
15	9	81.8	5	1 US-07-789-184-173	Sequence 173, App
16	9	81.8	5	1 US-08-406-192-12	Sequence 12, Appli
17	9	81.8	5	1 US-08-460-343B-37	Sequence 37, Appli
18	9	81.8	5	1 US-08-460-343B-38	Sequence 38, Appli
19	9	81.8	5	1 US-08-475-263-173	Sequence 173, App
20	9	81.8	5	1 US-08-398-028B-37	Sequence 37, Appli
21	9	81.8	5	1 US-08-398-028B-38	Sequence 38, Appli
22	9	81.8	5	1 US-08-266-407A-38	Sequence 38, Appli
23	9	81.8	5	1 US-08-485-866-173	Sequence 173, App
24	9	81.8	5	2 US-08-504-265B-37	Sequence 37, Appli
25	9	81.8	5	2 US-08-504-265B-38	Sequence 38, Appli
26	9	81.8	5	2 US-08-545-151-12	Sequence 12, Appli
27	9	81.8	5	2 US-08-392-973A-7	Sequence 7, Appli

28 9 81.8 5 2 US-08-392-973A-8 Sequence 8, Appli  
29 9 81.8 5 2 US-08-392-973A-9 Sequence 9, Appli  
30 9 81.8 5 2 US-08-477-362-173 Sequence 173, App  
31 9 81.8 5 2 US-08-477-134-173 Sequence 173, App  
32 9 81.8 5 2 US-08-892-544-38 Sequence 38, Appli  
33 9 81.8 5 3 US-08-473-489A-173 Sequence 173, App  
34 9 81.8 5 3 US-08-485-695-173 Sequence 173, App  
35 9 81.8 5 3 US-08-018-760-173 Sequence 173, App  
36 9 81.8 5 3 US-09-246-500B-17 Sequence 17, Appli  
37 9 81.8 5 4 US-09-367-791A-108 Sequence 108, App  
38 9 81.8 5 6 5169933-16 Patent No. 5169933  
39 9 81.8 5 6 5464756-42 Patent No. 5464756  
40 9 81.8 6 1 US-08-236-427-11 Sequence 11, Appli  
41 9 81.8 6 1 US-07-923-724-35 Sequence 35, Appli  
42 9 81.8 6 2 US-08-631-427A-3 Sequence 3, Appli  
43 9 81.8 6 2 US-08-806-203-2 Sequence 2, Appli  
44 9 81.8 6 2 US-08-806-203-3 Sequence 3, Appli  
45 9 81.8 6 2 US-08-806-203-4 Sequence 4, Appli

#### ALIGNMENTS

RESULT 1  
US-08-470-837-7  
; Sequence 7, Application US/08470837  
; Patent No. 5800811  
; GENERAL INFORMATION:  
; APPLICANT: Nimmi, Marcel E.  
; APPLICANT: Hall, Frederick L.  
; APPLICANT: Tuan, Tai-Jan  
; APPLICANT: Wu, Lingtao  
; APPLICANT: Cheung, David T.  
; TITLE OF INVENTION: Transforming Growth Factor B Fusion  
; TITLE OF INVENTION: and  
; TITLE OF INVENTION: Their Use in Wound Healing  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Merchand & Gould  
; STREET: 11150 Santa Monica Boulevard, Suite 400  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: USA  
; ZIP: 90025-3395  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/470,837  
; FILING DATE:  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharp, Janice A.  
; REGISTRATION NUMBER: 34,051  
; REFERENCE/DOCKET NUMBER: 30630-IUS01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 310-445-1140  
; TELEFAX: 310-445-9031  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-470-837-7

Query Match 81.8% Score 9; DB 1; Length 3;  
Best Local Similarity 66.7%; Pred No. 2.9e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
|  
Db 1 LTR 3

RESULT 2  
US-08-868-452-7  
; Sequence 7, Application US/08868452C  
; Patent No. 6352972  
; GENERAL INFORMATION:  
; APPLICANT: Marcel E. Nimmi  
; APPLICANT: Frederick L. Hall  
; APPLICANT: Lingtao Wu  
; APPLICANT: Bo Han  
; APPLICANT: Edwin Shors  
; TITLE OF INVENTION: BONE MORPHOGENETIC PROTEINS AND THEIR  
; FILE REFERENCE: 17972-11  
; CURRENT APPLICATION NUMBER: US/08/868,452C  
; CURRENT FILING DATE: 1997-06-03  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7  
; LENGTH: 3  
; TYPE: PRT  
; ORGANISM: Human  
US-08-868-452-7

Query Match 81.8%; Score 9; DB 4; Length 3;  
Best Local Similarity 66.7%; Pred. No. 2.9e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
|  
Db 1 LTR 3

RESULT 3  
US-08-288-405A-4  
; Sequence 4, Application US/08288405A  
; Patent No. 5559009  
; GENERAL INFORMATION:  
; APPLICANT: Chandy, Kanianthara G.  
; APPLICANT: Kalman, Katalin  
; APPLICANT: Chandy, Grischa  
; APPLICANT: Gutman, George A.  
; TITLE OF INVENTION: A No. 5559009el Voltage-Gated Potassium Channel  
; FILE REFERENCE: 10-AUG-1994  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert,  
; ADDRESSEE: Attn: Walter H. Dreger  
; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/288,405A  
; FILING DATE: 10-AUG-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/207,431  
; FILING DATE: 04-MAR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dreger, Walter H.  
; REGISTRATION NUMBER: 24,190  
; REFERENCE/DOCKET NUMBER: A-59844-1/WHD

TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 781-1989  
; TELEFAX: (415) 398-3249  
; TELEX: 910 277299  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-288-405A-4

Query Match 81.8%; Score 9; DB 1; Length 4;  
Best Local Similarity 66.7%; Pred. No. 2.9e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
|  
Db 2 LAR 4

RESULT 4  
US-08-336-343A-36  
; Sequence 36, Application US/08336343A  
; Patent No. 5677144  
; GENERAL INFORMATION:  
; APPLICANT: Ullrich, Axel  
; APPLICANT: Alves, Frauke  
; TITLE OF INVENTION: CCK-2, A No. 5677144el Receptor Tyrosine Kinase  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/336,343A  
; FILING DATE: 08-NOV-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7683-065  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-9741/8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-336-343A-36

Query Match 81.8%; Score 9; DB 1; Length 4;  
Best Local Similarity 66.7%; Pred. No. 2.9e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
|  
Db 2 LSR 4

RESULT 5  
US-09-142-078-19

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/533,889  
FILING DATE: 22 MAR-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/142,078  
FILING DATE: 10-FEB-1999  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO US97/12652  
FILING DATE: 21-JUL-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/762,377  
FILING DATE: 06-DEC-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/684,750

RESULT 6  
US-09-357-141-19  
; Sequence 19, Application US/09357141  
; Patent No. 6277825  
; GENERAL INFORMATION:  
; APPLICANT: Olivera, Baldomero W.  
; APPLICANT: McIntosh, J. Michael  
; APPLICANT: McCabe, R. Tyler  
; APPLICANT: Layer, Richard T.  
; APPLICANT: Zhou, Li-Ming

;; FILING DATE: 22-JUL-1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Ihnen, Jeffrey L.  
;; REGISTRATION NUMBER: 28,957  
;; REFERENCE/DOCKET NUMBER: 2314-168.A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 202-783-6040  
;; TELEFAX: 202-783-6031  
;; INFORMATION FOR SEQ ID NO: 19:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 4 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS:  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FRAGMENT TYPE: internal  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: 4  
;; OTHER INFORMATION: /note= "Xaa is  
;; gamma-carboxyglutamic acid"  
US-09-533-889-19  
Query Match 81.8%; Score 9; DB 4; Length 4;  
Best Local Similarity 66.7%; Pred. No. 2.9e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 LXR 3  
DB 1 LAR 3  
RESULT 8  
US-09-142-080-19  
; Sequence 19, Application US/09142080  
; Patent No. 6515103  
; GENERAL INFORMATION:  
; APPLICANT: Abogadie, Fe C.  
; Cruz, Lourdes J.  
; Olivera, Baldomero M.  
; Walker, Craig  
; Colledge, Clark  
; Hillyard, David R.  
; Jimenez, Elsie  
; Layer, Richard T.  
; Zhou, Li-Ming  
; McCabe, R. Tyler  
; TITLE OF INVENTION: Conantokins  
; NUMBER OF SEQUENCES: 71  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rothwell, Figg, Ernst & Manbeck, P.C.  
; STREET: 555 Thirteenth Street, N.W., Suite 701-E  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/142,080  
; FILING DATE: 11-May-2000  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO US97/12618  
; FILING DATE: 21-JUL-1997  
; APPLICATION NUMBER: US 08/684,742  
; FILING DATE: 22-JUL-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ihnen, Jeffrey L.  
; REGISTRATION NUMBER: 28,957  
; REFERENCE/DOCKET NUMBER: 2314-134.A

;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 202-783-6040  
;; TELEFAX: 202-783-6031  
;; INFORMATION FOR SEQ ID NO: 19:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 4 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: <Unknown>  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FRAGMENT TYPE: internal  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: 4  
;; OTHER INFORMATION: /note= "Xaa is  
;; gamma-carboxyglutamic acid"  
US-09-142-080-19  
Query Match 81.8%; Score 9; DB 4; Length 4;  
Best Local Similarity 66.7%; Pred. No. 2.9e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 LXR 3  
DB 1 LAR 3  
RESULT 9  
US-09-486-283C-4  
; Sequence 4, Application US/09486283C  
; Patent No. 6573243  
; GENERAL INFORMATION:  
; APPLICANT: ITAGAKI, Yasuhiro  
; APPLICANT: KONNO, Katsunori  
; APPLICANT: KAWAI, No. 6573243ufumi  
; APPLICANT: TAKAYAMA, Hiroaki  
; TITLE OF INVENTION: POMPIDID WASP-DERIVED NEUROPEPTIDES  
; FILE REFERENCE: 1830/48652  
; CURRENT APPLICATION NUMBER: US/09/486,283C  
; CURRENT FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: JP 9/241699  
; PRIOR FILING DATE: 1997-08-25  
; PRIOR APPLICATION NUMBER: PCT/JP98/03730  
; PRIOR FILING DATE: 1998-08-24  
; NUMBER OF SEQ ID NOS: 203  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 4  
; TYPE: PRT  
; ORGANISM: Pompidid sp.  
US-09-486-283C-4  
Query Match 81.8%; Score 9; DB 4; Length 4;  
Best Local Similarity 66.7%; Pred. No. 2.9e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 LXR 3  
DB 1 LSR 3  
RESULT 10  
US-09-535-852-355  
; Sequence 355, Application US/09535852  
; Patent No. 6638911  
; GENERAL INFORMATION:  
; APPLICANT: Blachuk, Orset W.  
; APPLICANT: Symonds, James M.  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND ETHNODS FOR MODULATING  
; TITLE OF INVENTION: DESMOSOMAL CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407C6

;; CURRENT APPLICATION NUMBER: US/09/535,852  
;; CURRENT FILING DATE: 2001-05-21  
;; NUMBER OF SEQ ID NOS: 2009  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 355  
;; LENGTH: 4  
;; TYPE: PRT  
;; ORGANISM: Unknown  
;; FEATURE:  
;; OTHER INFORMATION: No. 6638911classical cadherin calcium binding motif  
US-09-535-852-355

Query Match 81.8%; Score 9; DB 4; Length 4;  
Best Local Similarity 66.7%; Pred. No. 2.9e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
DB 1 LTR 3

RESULT 11  
US-07-626-923A-12  
;; Sequence 12, Application US/07626923A  
;; GENERAL INFORMATION:  
;; APPLICANT: Yoshimura, Akihiko  
;; APPLICANT: Longmore, Gregory D.  
;; APPLICANT: Louis, Harvey  
;; TITLE OF INVENTION: MUTANT EPO RECEPTOR AND USES  
;; TITLE OF INVENTION: THEREFOR  
;; NUMBER OF SEQUENCES: 14  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.  
;; STREET: Two Militia Drive  
;; CITY: Lexington  
;; STATE: Massachusetts  
;; COUNTRY: U.S.A.  
;; ZIP: 02173

COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/626,923A  
FILING DATE: 13 December 1990

CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: WH190-08  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 861-6240  
TELEFAX: (617) 861-9540  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-626-923A-12

Query Match 81.8%; Score 9; DB 1; Length 5;  
Best Local Similarity 66.7%; Pred. No. 2.9e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
DB 1 LAR 3

RESULT 12  
US-08-357-264-5

;; Sequence 5, Application US/08357264  
;; Patent No. 5541077  
;; GENERAL INFORMATION:  
;; APPLICANT: BURNIE Mr., James P.  
;; APPLICANT: MATTHEWS Ms., Ruth C.  
;; TITLE OF INVENTION: FUNGAL STRESS PROTEINS  
;; NUMBER OF SEQUENCES: 10  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Cushman, Darby & Cushman  
;; STREET: 1100 New York Avenue, N.W.  
;; CITY: Washington  
;; STATE: DC  
;; COUNTRY: USA  
;; ZIP: 20005-3918  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/357,264  
FILING DATE:

CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/152669  
FILING DATE: 16-NOV-1993  
APPLICATION NUMBER: US 663897  
FILING DATE: 14-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: KOKULIS, Paul N.  
REGISTRATION NUMBER: 16773  
REFERENCE/DOCKET NUMBER: PNK/5544/202253/DJP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-861-3000  
TELEFAX: 202-822-0944  
TELEX: 6714627 CUSH

INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-357-264-5

Query Match 81.8%; Score 9; DB 1; Length 5;  
Best Local Similarity 66.7%; Pred. No. 2.9e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
DB 1 LSR 3

RESULT 13  
US-08-456-840-38  
;; Sequence 38, Application US/08456840  
;; Patent No. 5597908  
;; GENERAL INFORMATION:  
;; APPLICANT: Taddei-Peters, W. C.  
;; APPLICANT: Butler, Sandra M.  
;; TITLE OF INVENTION: Immunoactive Peptides of Apo(a)  
;; NUMBER OF SEQUENCES: 48  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Akzo No. 5597908el  
;; STREET: 1330 Piccard Drive  
;; CITY: Rockville  
;; STATE: Maryland  
;; COUNTRY: US  
;; ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/456,840  
 FILING DATE: 01-JUN-1995  
 CLASSIFICATION: 536  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/266,407  
 FILING DATE: 27-JUN-1994  
 APPLICATION NUMBER: US 08/172,461  
 FILING DATE: 21-DEC-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Gortley, Mary E.  
 REGISTRATION NUMBER: 34,409  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 301-258-5200  
 TELEFAX: 301-977-0847  
 INFORMATION FOR SEQ ID NO: 38:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 5 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-456-840-38

Query Match 81.8%; Score 9; DB 1; Length 5;  
 Best Local Similarity 66.7%; Pred. No. 2.9e+05;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3  
 Db 1 LSR 3

RESULT 14  
 US-08-672-514-5  
 Sequence 5, Application US/08672514  
 Patent No. 5686248  
 GENERAL INFORMATION:  
 APPLICANT: BURNIE Mr., James P.  
 TITLE OF INVENTION: FUNGAL STRESS PROTEINS  
 NUMBER OF SEQUENCES: 10  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Cushman, Darby & Cushman  
 STREET: 1100 New York Avenue, N.W.  
 CITY: Washington  
 STATE: DC  
 COUNTRY: USA  
 ZIP: 20005-3918  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/672,514  
 FILING DATE: 28-JUN-1996  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/152,669  
 FILING DATE: 16-NOV-1993  
 APPLICATION NUMBER: US 08/152669  
 FILING DATE: 16-NOV-1993  
 APPLICATION NUMBER: US 663897  
 FILING DATE: 14-MAR-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: KOKULIS, Paul N.  
 REGISTRATION NUMBER: 16773  
 REFERENCE/DOCKET NUMBER: PNK/5544/202253/DJP  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-861-3000  
 TELEFAX: 202-822-0944  
 TELEX: 6714627 CUSH

INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 5 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-672-514-5

Query Match 81.8%; Score 9; DB 1; Length 5;  
 Best Local Similarity 66.7%; Pred. No. 2.9e+05;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3  
 Db 1 LSR 3

RESULT 15  
 US-07-789-184-173  
 Sequence 173, Application US/07789184  
 Patent No. 5688768  
 GENERAL INFORMATION:  
 APPLICANT: COUGHLIN, SHAUN R.  
 TITLE OF INVENTION: RECOMBINANT THROMBIN RECEPTOR AND  
 TITLE OF INVENTION: RELATED PHARMACEUTICALS  
 NUMBER OF SEQUENCES: 223  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: MORRISON & FOERSTER  
 STREET: 755 Page Mill Road  
 CITY: Palo Alto  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94304-1018  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA: US/07/789,184  
 APPLICATION NUMBER: 07/789,184  
 FILING DATE: 19911107  
 CLASSIFICATION: 514  
 ATTORNEY/AGENT INFORMATION:  
 NAME: MURASHIGE, KATE H.  
 REGISTRATION NUMBER: 29,959  
 REFERENCE/DOCKET NUMBER: 220000-20502.20  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 813-5600  
 TELEFAX: (415) 494-0792  
 TELEX: 34-0154  
 INFORMATION FOR SEQ ID NO: 173:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 5 amino acids  
 TYPE: AMINO ACID  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 FEATURE:  
 NAME/KEY: Modified-site  
 LOCATION: 5  
 OTHER INFORMATION: /note= "This position is R-NH2."  
 US-07-789-184-173

Query Match 81.8%; Score 9; DB 1; Length 5;  
 Best Local Similarity 66.7%; Pred. No. 2.9e+05;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3  
 Db 3 LAR 5

Job time : 3.48421 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 07:35:06 ; Search time 34.4421 Seconds  
(without alignments)  
37.345 Million cell updates/sec

Title: US-09-498-556C-357  
Perfect score: 11  
Sequence: 1 LXR 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 321558718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US05\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	81.8	4	9	US-09-248-158-2
2	9	81.8	4	9	US-09-818-656A-8
3	9	81.8	4	10	US-09-852-910-5
4	9	81.8	4	10	US-09-852-910-5
5	9	81.8	4	12	US-09-852-910-155
6	9	81.8	4	12	US-09-849-924-6
7	9	81.8	4	12	US-09-935-430-674
8	9	81.8	4	12	US-10-261-161-33
9	9	81.8	4	14	US-10-206-699-295
10	9	81.8	4	14	US-10-028-075B-6
11	9	81.8	4	14	US-10-029-206A-6
12	9	81.8	4	14	US-10-244-709-6
13	9	81.8	4	14	US-10-357-467-19
14	9	81.8	4	14	US-10-277-232-674
15	9	81.8	4	15	US-10-280-340-674
16	9	81.8	4	15	US-10-391-399-115

16	9	81.8	4	15	US-10-391-399-116	Sequence 116, Appl
17	9	81.8	4	15	US-10-411-336A-5	Sequence 5, Appl
18	9	81.8	4	15	US-10-411-336A-155	Sequence 155, Appl
19	9	81.8	5	12	US-10-243-613-80	Sequence 80, Appl
20	9	81.8	5	12	US-09-792-534A-7	Sequence 7, Appl
21	9	81.8	5	12	US-09-792-692B-7	Sequence 7, Appl
22	9	81.8	5	13	US-10-006-557-7	Sequence 7, Appl
23	9	81.8	5	14	US-10-206-699-32	Sequence 32, Appl
24	9	81.8	5	14	US-10-407-123-108	Sequence 108, Appl
25	9	81.8	5	14	US-10-303-817A-36	Sequence 36, Appl
26	9	81.8	5	15	US-10-394-980-334	Sequence 334, Appl
27	9	81.8	5	16	US-10-109-048-61	Sequence 61, Appl
28	9	81.8	6	9	US-09-486-734A-17	Sequence 17, Appl
29	9	81.8	6	9	US-09-486-734A-27	Sequence 27, Appl
30	9	81.8	6	9	US-09-770-102A-17	Sequence 17, Appl
31	9	81.8	6	10	US-09-866-512A-9	Sequence 9, Appl
32	9	81.8	6	10	US-09-500-700-73	Sequence 73, Appl
33	9	81.8	6	10	US-09-373-182B-1	Sequence 1, Appl
34	9	81.8	6	12	US-09-849-924-4	Sequence 4, Appl
35	9	81.8	6	12	US-09-935-430-693	Sequence 693, Appl
36	9	81.8	6	12	US-10-414-524-107	Sequence 107, Appl
37	9	81.8	6	12	US-10-328-953-52	Sequence 52, Appl
38	9	81.8	6	12	US-10-328-953-60	Sequence 60, Appl
39	9	81.8	6	12	US-10-367-580-322	Sequence 322, Appl
40	9	81.8	6	12	US-10-367-580-330	Sequence 330, Appl
41	9	81.8	6	12	US-10-367-593-322	Sequence 322, Appl
42	9	81.8	6	12	US-10-367-593-330	Sequence 330, Appl
43	9	81.8	6	12	US-10-367-594-322	Sequence 322, Appl
44	9	81.8	6	12	US-10-367-594-330	Sequence 330, Appl
45	9	81.8	6	12	US-10-367-654-322	Sequence 322, Appl

## ALIGNMENTS

RESULT 1  
US-09-248-158-2  
; Sequence 2, Application US/09248158  
; Patent No. US20020015678A1  
; GENERAL INFORMATION:  
; APPLICANT: Yuan, Zhengyu  
; APPLICANT: Chen, Zhong-Xiao  
; TITLE OF INVENTION: Direct Adsorption Scintillation Assay  
; TITLE OF INVENTION: For Measuring Enzyme Activity and Assaying Biochemical  
; TITLE OF INVENTION: Processes  
; FILE REFERENCE: 342312000600  
; CURRENT APPLICATION NUMBER: US/09/248,158  
; CURRENT FILING DATE: 1999-02-09  
; PRIOR APPLICATION NUMBER: 60/074,272  
; PRIOR FILING DATE: 1998-02-10  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 4  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Synthetic peptide  
US-09-248-158-2

Query Match	81.8%	Score 9;	DB 9;	Length 4;
Best Local Similarity	66.7%	Pred No. 1.2e+06;		
Matches	2;	Conservative	0;	Mismatches 1;
				Indels 0;
Gaps	0;			
Qy	1 LXR 3			
Db	2 LAR 4			
RESULT 2				
US-09-818-656A-8				
; Sequence 8, Application US/09818656A				
; Patent No. US20020142381A1				

```
; GENERAL INFORMATION:
; APPLICANT: GONG, Fangcheng et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: C0001191
; CURRENT APPLICATION NUMBER: US/09/818,656A
; CURRENT FILING DATE: 2000-03-28
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-09-818-656A-8

Query Match      81.8%; Score 9; DB 9; Length 4;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 2 LTR 4

RESULT 3
US-09-852-910-5
; Sequence 5, Application US/09852910
; Publication No. US20030096297A1
; GENERAL INFORMATION:
; APPLICANT: Hamm, Heidi
; TITLE OF INVENTION: Method For Identifying Inhibitors of G Protein Coupled Receptor S
; FILE REFERENCE: 2661-101
; CURRENT APPLICATION NUMBER: US/09/852,910
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US 60/275,472
; PRIOR FILING DATE: 2001-03-14
; NUMBER OF SEQ ID NOS: 271
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 4
; TYPE: PRT
; ORGANISM: mammal
; NAME/KEY: misc feature
; LOCATION: (1)..(4)
; OTHER INFORMATION: PAR-23
US-09-852-910-5

Query Match      81.8%; Score 9; DB 10; Length 4;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 1 LSR 3

RESULT 4
US-09-852-910-155
; Sequence 155, Application US/09852910
; Publication No. US20030096297A1
; GENERAL INFORMATION:
; APPLICANT: Hamm, Heidi
; TITLE OF INVENTION: Method For Identifying Inhibitors of G Protein Coupled Receptor S
; FILE REFERENCE: 2661-101
; CURRENT APPLICATION NUMBER: US/09/852,910
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US 60/275,472
; PRIOR FILING DATE: 2001-03-14
; NUMBER OF SEQ ID NOS: 271
```

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 155
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(4)
; OTHER INFORMATION: G alpha t library linker sequence
US-09-852-910-155

Query Match      81.8%; Score 9; DB 10; Length 4;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 2 LTR 4

RESULT 5
US-09-849-924-6
; Sequence 6, Application US/09849924
; Publication No. US20020037532A1
; GENERAL INFORMATION:
; APPLICANT: Regnier, Fred
; TITLE OF INVENTION: AFFINITY SELECTED SIGNATURE PEPTIDES FOR PROTEIN IDENTIFICATION A
; FILE REFERENCE: 290.00010101
; CURRENT APPLICATION NUMBER: US/09/849,924
; CURRENT FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: 60/203,227
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: 60/208,184
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 60/208,372
; PRIOR FILING DATE: 2000-05-31
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: C-terminal sequence
US-09-849-924-6

Query Match      81.8%; Score 9; DB 12; Length 4;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 2 LAR 4

RESULT 6
US-09-935-430-674
; Sequence 674, Application US/09935430
; Publication No. US20030017466A1
; GENERAL INFORMATION:
; APPLICANT: FARIS, MARY
; APPLICANT: HUBERT, RENE
; APPLICANT: RAITANO, ARTHUR
; APPLICANT: AFAR, DANIEL
; APPLICANT: LEVIN, ELANA
; APPLICANT: CHALLITA-EID, PIA
; APPLICANT: JAKOBOWITZ, AYA
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN NAMED 158PID7
; TITLE OF INVENTION: USEFUL IN THE TREATMENT AND DETECTION OF BLADDER AND
; TITLE OF INVENTION: OTHER CANCERS
; FILE REFERENCE: 51158-20050.00
; CURRENT APPLICATION NUMBER: US/09/935,430
```

```
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/227,098
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 700
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 674
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide motif
US-09-935-430-674

Query Match      81.8%; Score 9; DB 12; Length 4;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 LXR 3
Db      2 LTR 4

RESULT 7
US-10-261-161-33
; Sequence 33, Application US/10261161
; Publication No. US20040072270A1
; GENERAL INFORMATION:
; APPLICANT: Fernandez-Salas, Ester
; APPLICANT: Steward, Lance E.
; APPLICANT: Aoki, Kei Roger
; TITLE OF INVENTION: Cell-Based Fluorescence Resonance Energy
; TITLE OF INVENTION: Transfer (FRET) Assays For Clostridial Toxins
; FILE REFERENCE: P-AR 4804
; CURRENT APPLICATION NUMBER: US/10/261,161
; CURRENT FILING DATE: 2002-09-27
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
; NAME/KEY: REPEAT
; LOCATION: (1)...(4)
; OTHER INFORMATION: extent of repetition unknown.
US-10-261-161-33

Query Match      81.8%; Score 9; DB 12; Length 4;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 LXR 3
Db      1 LAR 3

RESULT 8
US-10-206-699-295
; Sequence 295, Application US/10206699
; Publication No. US20030100510A1
; GENERAL INFORMATION:
; APPLICANT: Sundaramoorthy, M.
; APPLICANT: Hudson, B.
; TITLE OF INVENTION: Crystallized structure of Type IV Collagen NC1 Domain Hexamer
; FILE REFERENCE: MBHB 01-1017
; CURRENT APPLICATION NUMBER: US/10/206,699
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/308,523
; PRIOR FILING DATE: 2001-07-27
```

```
; PRIOR APPLICATION NUMBER: US 60/351,289
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 60/366,854
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 60/385,362
; PRIOR FILING DATE: 2002-06-03
; NUMBER OF SEQ ID NOS: 307
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 295
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-206-699-295

Query Match      81.8%; Score 9; DB 14; Length 4;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 LXR 3
Db      1 LAR 3

RESULT 9
US-10-028-075B-6
; Sequence 6, Application US/10028075B
; Publication No. US20030113733A1
; GENERAL INFORMATION:
; APPLICANT: Khan, Nisar A.
; APPLICANT: Benner, Robert
; TITLE OF INVENTION: Gene regulator
; FILE REFERENCE: 2183-5223US
; CURRENT APPLICATION NUMBER: US/10/028,075B
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EP 01203748.7
; PRIOR FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 175
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: swiss/p36507/MPX2 Human
US-10-028-075B-6

Query Match      81.8%; Score 9; DB 14; Length 4;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 LXR 3
Db      2 LAR 4

RESULT 10
US-10-029-206A-6
; Sequence 6, Application US/10029206A
; Publication No. US20030119720A1
; GENERAL INFORMATION:
; APPLICANT: Khan, Nisar A.
; APPLICANT: Benner, Robert
; TITLE OF INVENTION: Oligopeptide treatment of anthrax
; FILE REFERENCE: 2183-5222US
; CURRENT APPLICATION NUMBER: US/10/029,206A
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/821,380
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 175
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 4
; TYPE: PRT
```

; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: peptide  
US-10-029-206A-6

Query Match 81.8%; Score 9; DB 14; Length 4;  
Best Local Similarity 66.7%; Pred. No. 1.2e+06;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3  
Db 2 LAR 4

RESULT 11  
US-10-244-709-6  
; Sequence 6, Application US/10244709  
; Publication No. US20030129769A1  
; GENERAL INFORMATION:  
; APPLICANT: FRED E. REGNIER  
; TITLE OF INVENTION: AFFINITY SELECTED SIGNATURE PEPTIDES FOR PROTEIN IDENTIFICATION A  
; FILE REFERENCE: 290.0001 0103  
; CURRENT APPLICATION NUMBER: US/10/244,709  
; CURRENT FILING DATE: 2002-09-16  
; PRIOR APPLICATION NUMBER: 60/203,227  
; PRIOR FILING DATE: 2000-05-05  
; PRIOR APPLICATION NUMBER: 60/208,184  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: 60/208,372  
; PRIOR FILING DATE: 2000-05-31  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 6  
; TYPE: PRT  
; LENGTH: 4  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: C-terminal sequence  
US-10-244-709-6

Query Match 81.8%; Score 9; DB 14; Length 4;  
Best Local Similarity 66.7%; Pred. No. 1.2e+06;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3  
Db 2 LAR 4

RESULT 12  
US-10-357-467-19  
; Sequence 19, Application US/10357467  
; Publication No. US20030194729A1  
; GENERAL INFORMATION:  
; APPLICANT: Abogadie, Fe C.  
; Cruz, Lourdes J.  
; Olivera, Baldomero M.  
; Walker, Craig  
; Colledge, Clark  
; Hillyard, David R.  
; Jimenez, Elsie  
; TITLE OF INVENTION: Conantokins  
; NUMBER OF SEQUENCES: 71  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rothwell, Figg, Ernst & Manbeck, p.c.  
; STREET: 1425 K Street, N.W., Suite 800  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/357,467  
; FILING DATE: 04-Feb-2003  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/142,080  
; FILING DATE: 15-MAY-2000  
; APPLICATION NUMBER: WO US97/12618  
; FILING DATE: 21-JUL-1997  
; APPLICATION NUMBER: US 08/684,742  
; FILING DATE: 22-JUL-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ihnen, Jeffrey L.  
; REGISTRATION NUMBER: 28,957  
; REFERENCE/DOCKET NUMBER: 2314-256.A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-783-6040  
; TELEFAX: 202-783-6031  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 4  
; OTHER INFORMATION: /note="Xaa is  
; gamma-carboxylglutamic acid"  
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-10-357-467-19

Query Match 81.8%; Score 9; DB 14; Length 4;  
Best Local Similarity 66.7%; Pred. No. 1.2e+06;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3  
Db 1 LAR 3

RESULT 13  
US-10-277-292-674  
; Sequence 674, Application US/10277292  
; Publication No. US20030199470A1  
; GENERAL INFORMATION:  
; APPLICANT: FARIS, MARY  
; APPLICANT: HUBERT, RENE  
; APPLICANT: RAITANO, ARTHUR  
; APPLICANT: AFAR, DANIEL  
; APPLICANT: LEVIN, ELANA  
; APPLICANT: CHALLITA-EID, PIA  
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN NAMED 158PID7  
; TITLE OF INVENTION: USEFUL IN THE TREATMENT AND DETECTION OF BLADDER AND  
; TITLE OF INVENTION: OTHER CANCERS  
; FILE REFERENCE: 51158-20050.00  
; CURRENT APPLICATION NUMBER: US/10/277,292  
; CURRENT FILING DATE: 2002-10-21  
; PRIOR APPLICATION NUMBER: US/09/935,430  
; PRIOR FILING DATE: 2001-08-22  
; PRIOR APPLICATION NUMBER: 60/227,098  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/282,739  
; PRIOR FILING DATE: 2001-04-10  
; NUMBER OF SEQ ID NOS: 700  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 674  
; LENGTH: 4

```
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide motif
US-10-277-292-674

Query Match      81.8%; Score 9; DB 14; Length 4;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 LXR 3
        |||
Db      2 LTR 4

RESULT 14
US-10-280-340-674
; Sequence 674, Application US/10280340
; Publication No. US20030207835A1
; GENERAL INFORMATION:
; APPLICANT: FARIS, MARY
; APPLICANT: HUBERT, RENE
; APPLICANT: RAITANO, ARTHUR
; APPLICANT: AFAR, DANIEL
; APPLICANT: LEVIN, ELANA
; APPLICANT: CHALLITA-EID, PIA
; APPLICANT: JAKOBOWITZ, AVA
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN NAMED 159P1D7
; TITLE OF INVENTION: USEFUL IN THE TREATMENT AND DETECTION OF BLADDER AND
; TITLE OF INVENTION: OTHER CANCERS
; FILE REFERENCE: 51158-20050.00
; CURRENT APPLICATION NUMBER: US/10/280,340
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: US/09/935,430
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/227,098
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 700
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 674
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide motif
US-10-280-340-674

Query Match      81.8%; Score 9; DB 15; Length 4;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 LXR 3
        |||
Db      2 LTR 4

RESULT 15
US-10-391-399-115
; Sequence 115, Application US/10391399
; Publication No. US20030219806A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Lora, Jose M.
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Silos-Santiago, Inmaculada
; TITLE OF INVENTION: NOVEL 18607, 15603, 69318, 12303, 48000,
; TITLE OF INVENTION: 52920, 38554, 57301, 58324, 55063,
; TITLE OF INVENTION: 52920, 38554, 57301, 58324, 55063,
; FILE REFERENCE: MPI03-0200NM7M
```

```
; CURRENT APPLICATION NUMBER: US/10/391,399
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 09/789,481
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: US 09/634,669
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: US 09/593,373
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: US 09/510,706
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 10/309,804
; PRIOR FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: US 60/336,936
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US 10/094,214
; PRIOR FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: US60/275,078
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 09/828,035
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/195,734
; PRIOR FILING DATE: 2000-04-07
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 115
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: N-glycosylation site
US-10-391-399-115

Query Match      81.8%; Score 9; DB 15; Length 4;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 LXR 3
        |||
Db      2 LSR 4

Search completed: September 24, 2004, 08:10:04
Job time : 34.4421 secs
```

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 07:28:33 ; Search time 2.44211 Seconds  
(without alignments)  
157.555 Million cell updates/sec

Title: US-09-498-556C-357

Perfect score: 11

Sequence: 1 LXR 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_78:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	81.8	7	E48394	glycoprotein compo
2	9	81.8	7	B48394	major fat-globule
3	9	81.8	8	PC4131	hypothetical prote
4	9	81.8	8	T13818	cytochrome oxidase
5	9	81.8	10	T13838	cytochrome-c oxida
6	9	81.8	10	T14219	cytochrome-c oxida
7	9	81.8	11	PU0034	dextranucrase (EC
8	9	81.8	12	A42324	cytochrome P450c27
9	9	81.8	13	PT0293	IG heavy chain CRD
10	9	81.8	14	PA0111	protein QA10054 -
11	9	81.8	15	S36896	ribosomal protein
12	9	81.8	15	PA0075	fructose-bisphosph
13	9	81.8	15	PA0102	fructose-bisphosph
14	9	81.8	15	B32800	hypothetical prote
15	9	81.8	15	PH1455	T-cell receptor al
16	9	81.8	15	AF0832	phe leader peptide
17	9	81.8	16	JH0517	insulin-like growt
18	9	81.8	16	C61414	chymotrypsin (EC 3
19	9	81.8	16	B44036	collagen alpha 1(X
20	9	81.8	17	JP0046	ribosomal protein
21	9	81.8	17	A34835	ribosomal protein
22	9	81.8	17	I54269	vitamin D binding
23	9	81.8	17	JQ2310	hypothetical 2-ik
24	9	81.8	17	JQ2320	hypothetical 2-ik
25	9	81.8	17	B61414	chymotrypsin (EC 3
26	9	81.8	18	S29491	GTP-binding protei
27	9	81.8	18	A41877	LcrKc - Versinia p
28	9	81.8	18	A43334	orf1 5' of aadr -
29	9	81.8	18	S33645	hypothetical prote

30 9 81.8 18 2 S10452 hypothetical prote  
31 9 81.8 19 2 S02269 glycogen(starch) s  
32 9 81.8 19 2 JC2060 homeobox 5 protein  
33 9 81.8 19 2 A60326 cholecystokinin-58  
34 9 81.8 20 2 A60525 lysozyme (EC 3.2.1  
35 9 81.8 20 2 JP0050 ribosomal protein  
36 9 81.8 20 2 JP0051 ribosomal protein  
37 9 81.8 20 2 S33001 hypothetical prote  
38 9 81.8 20 2 FC2084 serine proteinase  
39 9 81.8 20 2 S19616 globin - polychaet  
40 9 81.8 20 2 A61414 chymotrypsin (EC 3  
41 9 81.8 20 2 A42267 J-kappa recombinat  
42 9 81.8 20 2 AC0369 probable trp opero  
43 9 81.8 21 2 S16073 alanine-tRNA ligas  
44 9 81.8 22 2 JT0561 natriuretic peptid  
45 9 81.8 22 2 F23734 insulin-like growt

#### ALIGNMENTS

##### RESULT 1

E48394

glycoprotein component 16/major fat-globule membrane protein/MFG-E8 homolog - bovine (fr  
C;Species: Bos primigenius taurus (cattle)

C;Date: 19-Nov-1993 #sequence\_revision 18-Nov-1994 #text\_change 07-Feb-1997

C;Accession: E48394

R;Mather, I.H.; Banghart, L.R.; Lane, W.S.

Biochem. Mol. Biol. Int. 29, 545-554, 1993

A;Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-pig  
II-like sequences.

A;Reference number: A48394; MUID:93250576; PMID:8485470

A;Accession: E48394

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-7 <MAT>

A;Experimental source: milk

A;Note: sequence extracted from NCBI backbone (NCBIP:131450)

C;Keywords: glycoprotein

Query Match 81.8%; Score 9; DB 2; Length 7;  
Best Local Similarity 86.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 LXR 3

Db 5 LAR 7

##### RESULT 2

B48394

major fat-globule membrane protein GP 55 - guinea pig (fragment)

C;Species: Cavia porcellus (guinea pig)

C;Date: 19-Nov-1993 #sequence\_revision 18-Nov-1994 #text\_change 23-Mar-1995

C;Accession: E48394

R;Mather, I.H.; Banghart, L.R.; Lane, W.S.

Biochem. Mol. Biol. Int. 29, 545-554, 1993

A;Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-pig  
II-like sequences.

A;Reference number: A48394; MUID:93250576; PMID:8485470

A;Accession: E48394

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-7 <MAT>

A;Experimental source: milk

A;Note: sequence extracted from NCBI backbone (NCBIP:131444)

Query Match 81.8%; Score 9; DB 2; Length 7;  
Best Local Similarity 86.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 LXR 3

Db 5 LAR 7

## RESULT 3

PC4131

Hypothetical protein 8 [imported] - Pseudomonas aeruginosa (fragment)

C:Species: Pseudomonas aeruginosa

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: PC4131

R:Kawasaki, S.; Arai, H.; Igarashi, Y.; Kodama, T.

Gene 157, 87-91, 1995

A:Title: Sequencing and characterization of the downstream region of the genes encoding

Y for biosynthesis of heme d1.

A:Reference number: JC4552; MUID:96144254; PMID:8566817

A:Accession: PC4131

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-8 &lt;KAW&gt;

A:Cross-references: DDBJ:D50473; NID:gl1217594

A:Note: this ORF is not annotated in GenBank entry PSENIRC, release 113.0

Query Match

Best Local Similarity 81.8%; Score 9; DB 2; Length 8;

Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3

DB 4 LSR 6

## RESULT 4

T13818

Cytochrome oxidase subunit I - Atlantic hagfish mitochondrion (fragment)

C:Species: Mitochondrion Myxine glutinosa (Atlantic hagfish)

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jul-2000

C:Accession: T13818

R:Delabre, C.; Barriol, V.; Tillier, S.; Janvier, P.; Gachelin, G.

Mol. Biol. Evol. 14, 807-813, 1997

A:Title: The main features of the craniate mitochondrial DNA between the NDI and the COI

A:Reference number: Z17775; MUID:97398704; PMID:9254318

A:Accession: T13818

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-8 &lt;DEL&gt;

A:Cross-references: EMBL:Y09527; NID:g2340019; PIDN:CAA70718.1; PID:g2340022

A:Note: COI

A:Genome: mitochondrion

C:Keywords: mitochondrion

Query Match

Best Local Similarity 81.8%; Score 9; DB 2; Length 8;

Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3

DB 3 LSR 5

## RESULT 5

T13838

Cytochrome-c oxidase (EC 1.9.3.1) chain I - Bipes biporus mitochondrion (fragment)

C:Species: Mitochondrion Bipes biporus

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 11-May-2000

C:Accession: T13838

R:Macey, J.R.; Larson, A.; Ananjeva, N.B.; Fang, Z.; Papenfuss, T.J.

Mol. Biol. Evol. 14, 91-104, 1997

A:Title: Two novel gene orders and the role of light-strand replication in rearrangement

A:Reference number: Z17789; MUID:97153826; PMID:9000757

A:Accession: T13838

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-10 &lt;MAC&gt;

A:Cross-references: EMBL:U71335; NID:gl1753232; PID:gl1753235; PIDN:AAB48271.1

C:Genetics:

A:Genome: mitochondrion

A:Note: COI

C:Keywords: mitochondrion; oxidoreductase

Query Match

Best Local Similarity 81.8%; Score 9; DB 2; Length 10;

Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3

DB 3 LTR 5

## RESULT 6

T14219

Cytochrome-c oxidase (EC 1.9.3.1) chain I - Xenosaurus grandis mitochondrion (fragment)

C:Species: Mitochondrion Xenosaurus grandis

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jul-2000

C:Accession: T14219

R:Macey, J.R.; Larson, A.; Ananjeva, N.B.; Fang, Z.; Papenfuss, T.J.

Mol. Biol. Evol. 14, 91-104, 1997

A:Title: Two novel gene orders and the role of light-strand replication in rearrangement

A:Reference number: Z17789; MUID:97153826; PMID:9000757

A:Accession: T14219

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-10 &lt;MAC&gt;

A:Cross-references: EMBL:U71333; NID:G5739536; PIDN:AAC62821.1; PID:gl1753275

C:Genetics:

A:Genome: mitochondrion

A:Note: COI

C:Keywords: mitochondrion; oxidoreductase

Query Match

Best Local Similarity 81.8%; Score 9; DB 2; Length 10;

Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3

DB 3 LTR 5

## RESULT 7

PU0034

Dextranucrase (EC 2.4.1.5) - Streptococcus bovis (fragment)

C:Species: Streptococcus bovis

C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 18-Sep-1996

C:Accession: PU0034

R:Uezono, Y.; Tsumori, H.; Mukasa, H.

submitted to JIPID, October 1993

A:Description: Purification and properties of glucosyltransferase synthesizing 1,6-alpha-

A:Reference number: PU0034

A:Accession: PU0034

A:Molecule type: protein

A:Residues: 1-11 &lt;UEZ&gt;

A:Experimental source: ATCC 9809

C:Keywords: glycosyltransferase; hexosyltransferase

Query Match

Best Local Similarity 81.8%; Score 9; DB 2; Length 11;

Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3

DB 8 LTR 10

## RESULT 8

A42324

Cytochrome P450c27/25 - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C;Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 08-Feb-1996  
C;Accession: A42324  
R;Shayig, R.M.; Avadhani, N.G.  
J. Biol. Chem. 267, 2421-2428, 1992  
A;Title: Sequence complementarity between the 5'-terminal regions of mRNAs for rat mitochondria.  
A;Reference number: A42324; MUID:92129322; PMID:1733943  
A;Accession: A42324  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-12 <SHA>  
A;Note: sequence extracted from NCBI backbone (NCBIN:78404, NCBI:78410)

Query Match 81.8%; Score 9; DB 2; Length 12;  
Best Local Similarity 66.7%; Pred. No. 1.4e+03;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
|  
|  
|  
DB 4 LSR 6

RESULT 9  
PT0293  
Ig heavy chain CRD3 region (clone 4-124) - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
C;Accession: PT0293  
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
J. Exp. Med. 173, 395-407, 1991  
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and J<sub>H</sub> segments in the generation of the human Ig heavy chain repertoire.  
A;Reference number: PT0222; MUID:91108337; PMID:1893102  
A;Accession: PT0293  
A;Molecule type: DNA  
A;Residues: 1-13 <YAM>  
A;Experimental source: B lymphocyte  
A;Keywords: heterotrimer; immunoglobulin

Query Match 81.8%; Score 9; DB 2; Length 13;  
Best Local Similarity 66.7%; Pred. No. 1.5e+03;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
|  
|  
|  
DB 1 LSR 3

RESULT 10  
PA0111  
protein OA100054 - Arabidopsis thaliana (fragment)  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 07-Apr-1995 #sequence\_revision 26-May-1995 #text\_change 06-Jun-1997  
C;Accession: PA0111  
R;Kamo, M.; Kawakami, T.; Teugita, A.  
submitted to JIPID, March 1995  
A;Reference number: PA0109  
A;Accession: PA0111  
A;Molecule type: protein  
A;Residues: 1-14 <KAM>  
A;Experimental source: root

Query Match 81.8%; Score 9; DB 2; Length 14;  
Best Local Similarity 66.7%; Pred. No. 1.6e+03;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
|  
|  
|  
DB 11 LSR 13

RESULT 11  
S36896  
ribosomal protein S16 - Mycobacterium bovis (fragment)

C;Species: Mycobacterium bovis  
C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 13-Jan-1995  
C;Accession: S36896  
R;Ohara, N.; Kimura, M.; Higashi, Y.; Yamada, T.  
FEBS Lett. 331, 9-14, 1993  
A;Title: Isolation and amino acid sequence of the 30S ribosomal protein S19 from Mycobacterium bovis.  
A;Reference number: S36887; MUID:94009653; PMID:8405418  
A;Accession: S36896  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-15 <OHA>  
C;Keywords: protein biosynthesis; ribosome

Query Match 81.8%; Score 9; DB 2; Length 15;  
Best Local Similarity 66.7%; Pred. No. 1.7e+03;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
|  
|  
|  
DB 6 LTR 8

RESULT 12  
PA0075  
fructose-bisphosphate aldolase (EC 4.1.2.13) I - fungus (Fusarium sporotrichioides) (fragment)  
N;Alternate names: aldolase; fructose-1,6,-bisphosphate triosephosphate-lyase  
C;Species: Fusarium sporotrichioides  
C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 23-Mar-2001  
C;Accession: PA0075; PA0077  
R;Chow, L.P.; Fukaya, N.; Sugiura, Y.; Ueno, Y.; Tabuchi, K.; Teugita, A.  
submitted to JIPID, October 1994  
A;Description: Two dimensional polyacrylamide gel electrophoresis of Fusarium sporotrichioides.  
A;Reference number: PA0051  
A;Accession: PA0075  
A;Molecule type: protein  
A;Residues: 1-15 <CHO>  
A;Note: this form (II) had a molecular weight of 30.6K and an isoelectric point of 5.3  
A;Accession: PA0077  
A;Molecule type: protein  
A;Residues: 1-15 <CH2>  
A;Note: this form (III) had a molecular weight of 31.6K and an isoelectric point of 5.4  
C;Keywords: aldehyde-lyase; carbon-carbon lyase

Query Match 81.8%; Score 9; DB 2; Length 15;  
Best Local Similarity 66.7%; Pred. No. 1.7e+03;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
|  
|  
|  
DB 6 LSR 8

RESULT 13  
PA0102  
fructose-bisphosphate aldolase (EC 4.1.2.13) III - fungus (Fusarium sporotrichioides) (fragment)  
C;Species: Fusarium sporotrichioides  
C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 23-Mar-2001  
C;Accession: PA0102  
R;Chow, L.P.; Fukaya, N.; Sugiura, Y.; Ueno, Y.; Tabuchi, K.; Teugita, A.  
submitted to JIPID, October 1994  
A;Description: Two dimensional polyacrylamide gel electrophoresis of Fusarium sporotrichioides.  
A;Reference number: PA0051  
A;Accession: PA0102  
A;Molecule type: protein  
A;Residues: 1-15 <CHO>  
C;Keywords: aldehyde-lyase; carbon-carbon lyase

Query Match 81.8%; Score 9; DB 2; Length 15;  
Best Local Similarity 66.7%; Pred. No. 1.7e+03;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
|  
|  
|

Db 6 LSR 8

# RESULT 14

B32800  
 hypothetical protein (P1 5' region) - human (fragment)  
 C:Species: Homo sapiens (man)  
 C>Date: 22-Nov-1989 #sequence\_revision 22-Nov-1989 #text\_change 30-Sep-1993  
 C/Accession: B32800  
 R/Jindal, S.; Dudani, A.K.; Singh, B.; Harley, C.B.; Gupta, R.S.  
 Mol. Cell. Biol. 9, 2279-2283, 1989  
 A/Title: Primary structure of a human mitochondrial protein homologous to the bacterial  
 A/Reference number: A32800; MUID:89313783; PMID:2568584  
 A/Accession: B32800  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-15 <JIN>  
 A/Cross-references: GB:M23282

Query Match 81.8%; Score 9; DB 2; Length 15;  
 Best Local Similarity 66.7%; Pred. No. 1.7e+03;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3

Db 3 LSR 5

# RESULT 15

PH1455  
 T-cell receptor alpha chain (clone A24/PEF4) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C/Date: 10-Mar-1994 #sequence\_revision 10-Mar-1994 #text\_change 11-Apr-1995  
 C/Accession: PH1455  
 R/Casanova, J.L.; Martinon, F.; Gournier, H.; Barra, C.; Pannetier, C.; Regnault, A.; K  
 J. Exp. Med. 177, 811-820, 1993  
 A/Title: T cell receptor selection by and recognition of two class I major histocompatib  
 A/Reference number: PH1430; MUID:93171821; PMID:8436911  
 A/Accession: PH1455  
 A/Molecule type: mRNA  
 A/Residues: 1-15 <CAS>  
 A/Experimental source: cytolytic T-lymphocyte  
 C/Superfamily: immunoglobulin homology  
 C/Keywords: receptor; T-cell

Query Match 81.8%; Score 9; DB 2; Length 15;  
 Best Local Similarity 66.7%; Pred. No. 1.7e+03;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3

Db 3 LSR 5

Search completed: September 24, 2004, 07:36:05  
 Job time : 5.44211 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 07:28:33 ; Search time 1.38947 Seconds  
(without alignments)  
149.899 Million cell updates/sec

Title: US-09-498-556C-357

Perfect score: 11

Sequence: 1 LXR 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	81.8	11	RS30 ONCMY	P83328 oncorhynch
2	9	81.8	15	ONC1 ONCMY	P83287 oncorhynch
3	9	81.8	15	UC35 MAIZE	P80631 zea mays (m
4	9	81.8	16	IBP4 PIG	P24854 sus scrofa
5	9	81.8	18	YAA5 RHOPA	Q02005 rhodopsin
6	9	81.8	20	LYC FELCA	P37155 felis silve
7	9	81.8	20	RECK AZOVI	P37863 azotobacter
8	9	81.8	21	SYR RAT	P50475 rattus norv
9	9	81.8	22	ANFC CHICK	P21805 gallus gall
10	9	81.8	23	GLNA PHOLP	P20479 phormidium
11	9	81.8	25	IPYR PSEAN	P80898 pseudanaba
12	9	81.8	25	NEUT RANTE	P20056 rana tempor
13	9	81.8	25	SPIG PSEUS	P82357 pseudacanth
14	9	81.8	25	UBLI BOVIN	P23356 bos taurus
15	9	81.8	26	NTRC RHILP	P41502 rhizobium 1
16	9	81.8	26	YFHA KLEOX	P21710 klebsiella
17	9	81.8	27	CKXR CONRA	P58806 conus radia
18	9	81.8	27	SECR CANFA	P09910 canis fami
19	9	81.8	27	SECR RABIT	P32647 oryctolagus
20	9	81.8	27	SECS SHEEP	P31299 ovis aries
21	9	81.8	28	VI03 VACCP	Q00334 vaccinia vi
22	9	81.8	29	HS98 NEUCR	P31540 neurospora
23	9	81.8	29	Y51 BFT3	P20326 bacterioph
24	9	81.8	31	A98A DROME	O46201 drosophila
25	9	81.8	31	DIUX DIPPU	P82372 diptoptera
26	9	81.8	33	LYC2 HORSE	P81710 equus cabal
27	9	81.8	33	OTCC PSEPU	P11727 pseudomonas
28	9	81.8	33	PK1 DICDI	P34101 dictyosteli
29	9	81.8	34	RNLI PIG	P15466 sus scrofa
30	9	81.8	36	F4RE METOG	P80951 methanogeni
31	9	81.8	36	PAHO CERSI	P37999 ceratotheri
32	9	81.8	36	PAHO CHIBR	P41519 chinchilla
33	9	81.8	36	PAHO DIDMA	P18107 didelphis m

34 9 81.8 36 1 PAHO EQUZE P38000 equus zebra  
35 9 81.8 36 1 PAHO ERIEU P41335 erinaceus e  
36 9 81.8 36 1 PAHO MACMU P33684 macaca mula  
37 9 81.8 36 1 PAHO RABIT P41336 oryctolagus  
38 9 81.8 36 1 PAHO TAPPI P39659 tapirus pin  
39 9 81.8 36 1 PYV AMICA P29205 amia calva  
40 9 81.8 36 1 PYV PIG P01305 sus scrofa  
41 9 81.8 36 1 YHDS BACSU O07588 bacillus su  
42 9 81.8 37 1 CALI PIG P30880 sus scrofa  
43 9 81.8 37 1 CALI SHEEP P30881 ovis aries  
44 9 81.8 37 1 CALR RANRI P31988 rana ridibu  
45 9 81.8 37 1 PIP7 BOVIN P21671 bos taurus

#### ALIGNMENTS

RESULT 1  
RS30 ONCMY STANDARD; PRT; 11 AA.  
ID P83328; 2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE 40S ribosomal protein S30 (Fragment).  
GN PAU.  
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
OX NCBI\_TaxID=8022;  
[1]  
RN RAINBOW TROUT  
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.  
EC TISSUE=Skin mucus;  
RX MEDLINE=22142142; PubMed=12147245;  
RA Fernandes J.M.O., Smith V.J.;  
RT "A novel antimicrobial function for a ribosomal peptide from rainbow trout skin."  
RL Biochem. Biophys. Res. Commun. 296:167-171(2002).  
CC -!- FUNCTION: Has antibacterial activity against Gram-positive bacteria.  
CC -!- MASS SPECTROMETRY: MW=6676.6; METHOD=MALDI.  
CC -!- SIMILARITY: Belongs to the S30E family of ribosomal proteins.  
KW Ribosomal protein; Antibiotic.  
FT NON TER 11  
SQ SEQUENCE 11 AA; 1123 MW; 2312AB630DD735B8 CRC64;  
Query Match 81.8%; Score 9; DB 1; Length 11;  
Best Local Similarity 86.7%; Pred. No. 4e+02; 1; Indels 0; Gaps 0;  
Matches 2; Conservative 0; Mismatches 1;  
Qy 1 LXR 3  
Db 6 LAR 8  
RESULT 2  
ONC1 ONCMY STANDARD; PRT; 15 AA.  
ID P83287; 2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
OX NCBI\_TaxID=8022;  
[1]  
RN RAINBOW TROUT  
RP SEQUENCE, FUNCTION, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.  
EC TISSUE=Skin mucus;  
RX MEDLINE=20394650; PubMed=10938737;

```

RA Smith V.J., Fernandes J.M.O., Jones S.J., Kemp G.D., Tatner M.F.;
RT "Antibacterial proteins in rainbow trout, Oncorhynchus mykiss.";
RL Fish Shellfish Immunol. 10:243-260(2000).
CC -1- FUNCTION: Has antibacterial activity against Gram-positive
CC bacterium P.citreus.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Skin.
DR GO: GO:0005576; C:extracellular; NAS.
DR GO: GO:0003795; F:antimicrobial peptide activity; NAS.
DR GO: GO:0006805; P:xenobiotic metabolism; NAS.
KW Antibiotic.
FT UNSURE 4 4 OR G.
FT UNSURE 9 9 OR T.
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1601 MW; 43C25028DEC12B7C CRC64;

Query Match 81.8%; Score 9; DB 1; Length 15;
Best Local Similarity 66.7%; Pred. No. 5.7e+02;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
Db 12 LAR 14

RESULT 3
UC25-MAIZE STANDARD; PRT; 15 AA.
ID UC25-MAIZE STANDARD; PRT; 15 AA.
AC P80631;
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Unknown protein from 2D-page of etiolated coleoptile (Spot 77)
DE (Fragment).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE.
RC TISSUE=Coleoptile;
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA Pernollet J.-C., Zivy M., de Vienne D.;
RT "The maize two dimensional gel protein database: towards an integrated
RT genome analysis program.";
RL Theor. Appl. Genet. 93:957-1005(1996).
CC -1- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
CC protein is: 4.9, its MW is: 31.6 kDa.
DR Maize-2DPAGE; P80631; COLEOPTILE.
DR MaizedB; 123957; -.
FT NON_TER 1 1
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1580 MW; 83C54CF0CE1614D0 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 15;
Best Local Similarity 66.7%; Pred. No. 5.7e+02;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
Db 5 LSR 7

RESULT 4
IBP4-PIG STANDARD; PRT; 16 AA.
ID IBP4-PIG STANDARD; PRT; 16 AA.
AC P24854;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Insulin-like growth factor binding protein 4 (IGFBP-4) (IBP-4)
DE (IGF-binding protein 4) (Fragment).

```

```

GN IGFBP4.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE.
RX MEDLINE=92109718; PubMed=1722398;
RA Coleman M.E., Pan Y.-C.E., Etherton T.D.;
RT "Identification and NH2-terminal amino acid sequence of three
RT insulin-like growth factor-binding proteins in porcine serum.";
RL Biochem. Biophys. Res. Commun. 181:1131-1136(1991).
CC -1- FUNCTION: IGF-binding proteins prolong the half-life of the IGFs
CC and have been shown to either inhibit or stimulate the growth
CC promoting effects of the IGFs on cell culture. They alter the
CC interaction of IGFs with their cell surface receptors.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 thyroglobulin type-I domain.
DR PIR; JH0517; JH0517.
DR InterPro; IPR000867; Insl_gro_fac_pr.
DR InterPro; IPR00716; Thyroglobulin_1.
DR PROSITE; PS00222; IGF_BINDING; PARTIAL.
DR PROSITE; PS00484; THYROGLOBULIN_1; PARTIAL.
KW Growth factor binding.
FT NON_TER 16 16
SQ SEQUENCE 16 AA; 1799 MW; 40988840096655E2 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 16;
Best Local Similarity 56.7%; Pred. No. 6.1e+02;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
Db 14 LAR 16

RESULT 5
YAAS-RHOA STANDARD; PRT; 18 AA.
ID YAAS-RHOA STANDARD; PRT; 18 AA.
AC Q02005;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein in aaDr 5' region (Fragment).
OS Rhodopseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodospseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CGA009;
RX MEDLINE=92394882; PubMed=1522059;
RA Dispensa M., Thomas C.T., Kim M.K., Perrotta J.A., Gibson J.,
RA Harwood C.S.;
RT "Anaerobic growth of Rhodopseudomonas palustris on 4-hydroxybenzoate
RT is dependent on AaDr, a member of the cyclic AMP receptor protein
RT family of transcriptional regulators.";
RL J. Bacteriol. 174:5803-5813(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M92426; AAA26089.1; -.
DR PIR; A43334; A43334.
KW Hypothetical protein.
FT NON_TER 1 1
SQ SEQUENCE 18 AA; 2173 MW; 61DABDFE4E317B0A CRC64;

```

Query Match 81.8%; Score 9; DB 1; Length 18;  
 Best Local Similarity 66.7%; Pred. No. 6.9e+02;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
 |  
 |  
 6 LTR 8

DB

RESULT 6  
 LYC\_FELCA  
 ID LYC\_FELCA STANDARD; PRT; 20 AA.  
 AC P37155;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Lysozyme C (EC 3.2.1.17) (1,4-beta-N-acetylmuramidase C) (Fragment).  
 GN LYZ.

OS Felis silvestris catus (Cat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
 OC NCBI\_TaxID=9685;  
 RN [1]  
 RP SEQUENCE  
 RC TISSUE=Milk;  
 RX MEDLINE=90263403; PubMed=2344734;  
 RA Halliday J.A., Bell K., McKenzie H.A., Shaw D.C.;  
 RT "Feline whey proteins: identification, isolation and initial  
 RT characterization of alpha-lactalbumin, beta-lactoglobulin and  
 RT lysozyme";  
 RL Comp. Biochem. Physiol. 95B:773-779(1990).  
 CC -!- FUNCTION: Lysozymes have primarily a bacteriolytic function; those  
 CC in tissues and body fluids are associated with the monocyte-  
 CC macrophage system and enhance the activity of immunogens.  
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages between N-  
 CC acetyl-D-glucosamine and N-acetylmuramic acid in peptidoglycan  
 CC heteropolymers of the prokaryotes cell walls.  
 CC -!- SUBUNIT: Monomer.  
 CC -!- MISCELLANEOUS: Lysozyme C is capable of both hydrolysis and  
 CC transglycosylation; it shows also a slight esterase activity. It  
 CC acts rapidly on both peptide-substituted and unsubstituted  
 CC peptidoglycan, and slowly on chitin oligosaccharides.  
 CC -!- SIMILARITY: Belongs to family 22 of glycosyl hydrolases.  
 DR PIR; A60525; A60525.  
 DR HSSP; P11376; 2EQI.  
 DR InterPro; IPR001916; Glyco\_hydro\_22.  
 DR Pfam; PF00062; Iys; 1.  
 DR PROSITE; PS00128; LACTALBUMIN\_LYSOZYME; PARTIAL.  
 KW Hydrolase; Glycosidase; Bacteriolytic enzyme; Milk.  
 FT NON TER 20  
 SQ SEQUENCE 20 AA; 2314 MW; EB8824EA59425E13 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 20;  
 Best Local Similarity 66.7%; Pred. No. 7.7e+02;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
 |  
 |  
 8 LAR 10

DB

RESULT 7  
 RECX\_AZOVI  
 ID RECX\_AZOVI STANDARD; PRT; 20 AA.  
 AC P37863;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Regulatory protein reCX (Fragment).  
 GN RECX.  
 OS Azotobacter vinelandii.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

Pseudomonadaceae; Azotobacter.  
 OC NCBI\_TaxID=354;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92225347; PubMed=1563632;  
 RA Venkatesh T.V., Das H.K.;  
 RT "The Azotobacter vinelandii recA gene: sequence analysis and  
 RT regulation of expression";  
 RL Gene 113:47-53(1992).  
 RN [2]  
 RP IDENTIFICATION.  
 RX MEDLINE=94218258; PubMed=8165147;  
 RA de Mot R., Schoofs G., Vanderleyden J.;  
 RT "A putative regulatory gene downstream of recA is conserved in Gram-  
 RT negative and Gram-positive bacteria";  
 RL Nucleic Acids Res. 22:1313-1314(1994).  
 CC -!- FUNCTION: Modulates recA activity (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -!- SIMILARITY: Belongs to the recX family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; S96898; -; NOT ANNOTATED\_CDS.  
 DR HAMAP; MF\_01114; -; 1.  
 FT NON TER 20  
 SQ SEQUENCE 20 AA; 2111 MW; C809F8BCCED6CB56 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 20;  
 Best Local Similarity 66.7%; Pred. No. 7.7e+02;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
 |  
 |  
 18 LAR 20

DB

RESULT 8  
 SYA\_RAT  
 ID SYA\_RAT STANDARD; PRT; 21 AA.  
 AC P50475;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Alanyl-tRNA synthetase (EC 6.1.1.7) (Alanine--tRNA ligase) (AlaRS)  
 DE (Fragment).  
 GN AARS.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OC NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE.  
 RX TISSUE=Liver;  
 RX MEDLINE=91249799; PubMed=2040280;  
 RA Dignam J.D., Dignam S.S., Brumley L.L.;  
 RT "Alanyl-tRNA synthetase from Escherichia coli, Bombyx mori and Ratus  
 RT ratus. Existence of common structural features";  
 RL Eur. J. Biochem. 198:201-210(1991).  
 CC -!- CATALYTIC ACTIVITY: ATP + L-alanine + tRNA(Ala) = AMP +  
 CC diphosphate + L-alanyl-tRNA(Ala).  
 CC -!- SUBUNIT: Monomer.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.  
 CC PIR; S16073; S16073.  
 DR InterPro; IPR006193; tRNA\_synth\_Ala.  
 DR PROSITE; PS50860; AA\_TRNA\_LIGASE\_II\_ALA; PARTIAL.  
 DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.

FT NON\_TER 1 1  
 FT NON\_TER 21 21  
 SQ SEQUENCE 21 AA; 2293 MW; D739DDC62CD43375 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 21;  
 Best Local Similarity 66.7%; Pred. No. 8.2e+02;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
 |  
 |  
 Db 18 LAR 20

## RESULT 9

ANFC\_CHICK STANDARD; PRT; 22 AA.  
 AC P21805;  
 DT 01-MAY-1991 (Rel. 18, Created)  
 DT 01-MAY-1991 (Rel. 18, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE C-type natriuretic peptide (CNP).  
 GN NPFC.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]

## SEQUENCE.

RP TISSUE=Brain;  
 RX MEDLINE=91113186; PubMed=1989595;  
 RA Arimura J.J., Minamino N., Kangawa K., Matsuo H.;  
 RT "Isolation and identification of C-type natriuretic peptide in  
 chicken brain";  
 RL Biochem. Biophys. Res. Commun. 174:142-148(1991).  
 CC -!- FUNCTION: Vaso-relaxant activity. Has a cGMP-stimulating activity.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the natriuretic peptide family.  
 DR PIR; JT0581; JT0581.  
 DR InterPro; IPR000663; Natr\_peptide.  
 DR Pfam; PF00212; ANP; 1.  
 DR PRINTS; PR00710; NATPEPTIDES.  
 DR SMART; SM00183; NAT\_PEP; 1.  
 DR PROSITE; PS00263; NATRIURETIC\_PEPTIDE; 1.  
 FT DISULFID 6 22  
 SQ SEQUENCE 22 AA; 2244 MW; FC2A4706D8DAC025 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 22;  
 Best Local Similarity 66.7%; Pred. No. 8.6e+02;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
 |  
 |  
 Db 2 LSR 4

## RESULT 10

GLNA\_PHOLP STANDARD; PRT; 23 AA.  
 AC P20479;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Glutamine synthetase (EC 6.3.1.2) (Glutamate--ammonia ligase)  
 DE (Fragment).  
 OS Phormidium lapideum.  
 OC Bacteria; Cyanobacteria; Oscillatoriales; Phormidium.  
 OX NCBI\_TaxID=32060;  
 RN [1]

## SEQUENCE.

RP MEDLINE=89214011; PubMed=2907514;  
 RX Sawa Y., Ochiai H., Yoshida K., Tanizawa K., Tanaka H., Soda K.;  
 RA

RT "Glutamine synthetase from a cyanobacterium, Phormidium lapideum:  
 purification, characterization, and comparison with other  
 cyanobacterial enzymes.";  
 RL J. Biochem. 104:917-923(1988).  
 CC -!- CATALYTIC ACTIVITY: ATP + L-glutamate + NH(3) = ADP + phosphate +  
 L-glutamine.  
 CC -!- SUBUNIT: Oligomer of 12 subunits arranged in the form of two  
 hexagons.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: Belongs to the glutamine synthetase family.

DR PIR; PX0011; PX0011.  
 DR InterPro; IPR008147; Gln\_synt\_beta.  
 DR InterPro; IPR008146; Gln\_synt\_C.  
 DR PROSITE; PS00180; GLNA\_1; PARTIAL.  
 DR PROSITE; PS00181; GLNA\_ATP; PARTIAL.  
 KW Ligase.  
 FT NON\_TER 23 23

SQ SEQUENCE 23 AA; 2656 MW; 20B69C164D2A5739 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 23;  
 Best Local Similarity 66.7%; Pred. No. 9e+02;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
 |  
 |  
 Db 7 LSR 9

## RESULT 11

IPYR\_PSEAN STANDARD; PRT; 25 AA.  
 ID IPYR\_PSEAN  
 AC P80838;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Inorganic pyrophosphatase (EC 3.6.1.1) (Pyrophosphate phospho-  
 hydrolyase) (PPase) (Fragment).  
 GN PPA.  
 OS Pseudanabaena sp. (strain PCC 6901).  
 OC Bacteria; Cyanobacteria; Oscillatoriales;  
 OC Pseudanabaena/limnithrix group; Pseudanabaena.  
 OX NCBI\_TaxID=47918;  
 RN [1]

## SEQUENCE.

RA Gomez R., Serrano A.;  
 RL Submitted (FEB-1997) to Swiss-Prot.

CC -!- FUNCTION: Hydrolyzes PPI generated in anabolic reactions.

CC -!- CATALYTIC ACTIVITY: Diphosphate + H(2)O = 2 phosphate.  
 CC -!- COFACTOR: Binds 4 magnesium ions per subunit. Other metal ions can  
 support activity, but at a lower rate. Two magnesium ions are  
 required for the activation of the enzyme and are present before  
 substrate binds, two additional magnesium ions form complexes with  
 substrate and product (3y similarity).

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: Belongs to the PPase family.

DR HAMAP; MF\_00209; -; 1.

KW Hydrolyase; Metal-binding; Magnesium.

FT NON\_TER 25 25

SQ SEQUENCE 25 AA; 2630 MW; FB6138F3AED4D43E CRC64;

Query Match 81.8%; Score 9; DB 1; Length 25;  
 Best Local Similarity 66.7%; Pred. No. 9.9e+02;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
 |  
 |  
 Db 3 LSR 5

## RESULT 12

NEUU\_RANTE STANDARD; PRT; 25 AA.  
 ID NEUU\_RANTE  
 AC P20056;

DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Neuromedin U-25 (NMU-25).  
 OS Rana temporaria (European common frog).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Amphibia; Batrachia; Anura; Neobatrachia; Ranidae; Rana.  
 OX NCBI\_TaxID=8407;  
 RN [1]  
 RN SEQUENCE.  
 RC TISSUE=Intestine;  
 RX MEDLINE=90078173; PubMed=2592357;  
 RA Domin J., Yiangou Y.G., Spokes R.A., Aitken A., Parmar K.B.,  
 RA Chrysanthou B.J., Bloom S.R.;  
 RT "The distribution, purification, and pharmacological action of an  
 RT amphibian neuromedin U.";  
 RL J. Biol. Chem. 264:20881-20885(1989).  
 CC -!- FUNCTION: Stimulates uterine smooth muscle contraction and causes  
 CC selective vasoconstriction.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the NMU family.  
 DR PIR; A34179; A34179.  
 DR InterPro; IPR008199; NMU.  
 DR SMART; SM00084; NMU; 1.  
 DR PROSITE; PS00967; NMU; 1.  
 KW Amidation; Hormone.  
 FT MOD\_RES 25  
 FT SEQUENCE 25 AA; 2632 MW; 6A01D89F6DA06FD4 CRC64;  
 SQ SEQUENCE 25 AA; 2632 MW; 6A01D89F6DA06FD4 CRC64;  
 Query Match 81.8%; Score 9; DB 1; Length 25;  
 Best Local Similarity 66.7%; Pred. No. 9.9e+02;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 LXR 3  
 Db 14 LSR 16

RESULT 13  
 SP1G\_PSEUS  
 ID SP1G\_PSEUS STANDARD; PRT; 25 AA.  
 AC P82357;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Spingerin.  
 OS Pseudacanthotermes spingeri.  
 CC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;  
 CC Neoptera; Orthopteroidea; Dictyoptera; Isoptera; Termitidae;  
 CC Macrotermittinae; Pseudacanthotermes.  
 OX NCBI\_TaxID=115113;  
 RN [1]  
 RN SEQUENCE, MASS SPECTROMETRY, AND FUNCTION.  
 RC TISSUE=Blood, and Salivary gland;  
 RX PubMed=11053427;  
 RA Lamberty M., Zachary D., Bordenau C., Robert A.,  
 RA Hoffmann J.A., Bulet P.;  
 RT "Insect Immunity. Constitutive expression of a cysteine-rich  
 RT antifungal and a linear antibacterial peptide in a termite insect.";  
 RL J. Biol. Chem. 276:4085-4092(2001).  
 CC -!- FUNCTION: Active against Gram-positive bacteria B.megaterium and  
 CC M.luteus, Gram-negative bacteria E.coli SBS363 and D22,  
 CC K.pneumoniae, S.typhimurium and P.aeruginosa, yeast C.albicans  
 CC and filamentous fungi F.culmorum, N.crassa, N.hematococca and  
 CC T.vividae. Inactive against Gram-positive bacteria B.subtilis,  
 CC S.pyogenes, B.thuringiensis and S.aureus, Gram-negative bacteria  
 CC E.coliaceae and E.carotovora and filamentous fungus B.bassiana.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- INDUCTION: By bacterial infection.  
 CC -!- MASS SPECTROMETRY: MW=3001.8; METHOD=MALDI.  
 CC -!- MISCELLANEOUS: There are three isoforms of spingerin.  
 KW Antibiotic; Fungicide.

FT VARIANT 1 3 MISSING (IN N-3 ISOFORM).  
 FT VARIANT 22 25 MISSING (IN C-4 ISOFORM).  
 SQ SEQUENCE 25 AA; 3001 MW; AA79370264262F60 CRC64;  
 Query Match 81.8%; Score 9; DB 1; Length 25;  
 Best Local Similarity 66.7%; Pred. No. 9.9e+02;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 LXR 3  
 Db 22 LTR 24  
 RESULT 14  
 UBLL\_BOVIN  
 ID UBLL\_BOVIN STANDARD; PRT; 25 AA.  
 AC P23356;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Ubiquitin carboxyl-terminal hydrolase isozyme L1 (EC 3.4.19.12) (UCH-  
 DE L1) (Ubiquitin thiolesterase L1) (Neuron cytoplasmic protein 9.5)  
 DE (PGP 9.5) (PGP9.5) (Fragment).  
 GN UCHL1.  
 OS Bos taurus (Bovine).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 CC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RN SEQUENCE.  
 RC TISSUE=Brain;  
 EX MEDLINE=92008646; PubMed=1833240;  
 RA Giambanco I., Bianchi R., Ceccarelli P., Pula G., Sorci G.,  
 RA Antonelli S., Bocchini V., Donato R.;  
 RT "Neuron-specific" protein gene product 9.5 (PGP 9.5) is also  
 RT expressed in glioma cell lines and its expression depends on cellular  
 RT growth state.";  
 RL FEBS Lett. 290:131-134(1991).  
 CC -!- FUNCTION: Ubiquitin-protein hydrolase is involved both in the  
 CC processing of ubiquitin precursors and of ubiquitinated proteins.  
 CC This enzyme is a thiol protease that recognizes and hydrolyzes  
 CC a peptide bond at the C-terminal glycine of ubiquitin.  
 CC -!- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =  
 CC ubiquitin + a thiol.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- TISSUE SPECIFICITY: Neurons and cells of the diffuse  
 CC neuroendocrine system and their tumors.  
 CC -!- SIMILARITY: Belongs to peptidase family C12.  
 DR PIR; S17561; S17561.  
 DR MEROPS; C12.001; -.  
 DR InterPro; IPR001578; Peptidase C12.  
 DR Pfam; PF01088; Peptidase C12; 1.  
 DR PROSITE; PS00140; UCH1; PARTIAL.  
 KW Ub1 conjugation pathway; Hydrolase; Thiol protease; Multigene family.  
 FT NON\_TER 25  
 FT SEQUENCE 25 AA; 2812 MW; 26BB5ADD0A754D55 CRC64;  
 Query Match 81.8%; Score 9; DB 1; Length 25;  
 Best Local Similarity 66.7%; Pred. No. 9.9e+02;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 LXR 3  
 Db 17 LTR 19  
 RESULT 15  
 NTRC\_RHLP  
 ID NTRC\_RHLP STANDARD; PRT; 26 AA.  
 AC P41502;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Nitrogen assimilation regulatory protein (fragment).  
 GN NTRC.  
 OS Rhizobium leguminosarum (biovar phaseoli).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.  
 OX NCBI\_TaxID=385;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CE-3;  
 RX MEDLINE=94018651; PubMed=8412703;  
 RA Patriarca E.J., Riccio A., Tate R., Colonna-Romano S., Iaccarino M.,  
 RA Defez R.;  
 RT "The ntrBC genes of Rhizobium leguminosarum are part of a complex  
 RT operon subject to negative regulation.";  
 RL Mol. Microbiol. 9:569-577(1993).  
 CC -1- FUNCTION: MEMBER OF THE TWO-COMPONENT REGULATORY SYSTEM NTRB/NTRC  
 CC INVOLVED IN THE ACTIVATION OF NITROGEN ASSIMILATORY GENES SUCH AS  
 CC GlnA. NTRC IS PHOSPHORYLATED BY NTRB AND INTERACTS WITH SIGMA-54.  
 CC -1- SIMILARITY: Contains 1 response regulatory domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X71436; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; S36203; S36203.  
 DR InterPro; IPR001789; Response\_reg.  
 DR PROSITE; PS50110; RESPONSE REGULATORY; 1.  
 KW Nitrogen fixation; Transcription regulation; Repressor; Activator;  
 KW DNA-binding; ATP-binding; Phosphorylation; Sensory transduction.  
 FT DOMAIN 1 >26 RESPONSE REGULATORY.  
 FT NON\_TER 26  
 SQ SEQUENCE 26 AA; 2687 MW; E11ECC63C8636304 CRC64;  
  
 Query Match 81.8%; Score 9; DB 1; Length 26;  
 Best Local Similarity 66.7%; Pred. No. 1e+03;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 1 LXR 3  
 Db |  
 22 LSR 24

Search completed: September 24, 2004, 07:35:01  
 Job time : 4.38947 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 07:28:32 ; Search time 4.16842 Seconds  
(without alignments)  
302.770 Million cell updates/sec

Title: US-09-498-556C-357  
Perfect score: 11  
Sequence: 1 LXR 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 25:.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mbc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriaph.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	81.8	8	Q16468	Q16468 homo sapien
2	9	81.8	8	Q974Y2	Q974Y2 asterina pe
3	9	81.8	8	Q89965	Q89965 polyomaviru
4	9	81.8	9	Q91688	Q91688 gekko gekko
5	9	81.8	9	Q94VH4	Q94VH4 varanus gla
6	9	81.8	9	Q94VD8	Q94VD8 varanus nil
7	9	81.8	9	Q94V18	Q94V18 varanus ere
8	9	81.8	9	Q94VC6	Q94VC6 varanus pil
9	9	81.8	9	Q94VE1	Q94VE1 varanus mer
10	9	81.8	9	Q91BMS	Q91BMS simian viru
11	9	81.8	9	Q91YKL	Q91YKL simian viru
12	9	81.8	10	Q958K9	Q958K9 rana boylii
13	9	81.8	10	Q97G86	Q97G86 diploglossu
14	9	81.8	10	Q94V97	Q94V97 varanus spe
15	9	81.8	10	Q94VD5	Q94VD5 varanus oli
16	9	81.8	10	Q94VC9	Q94VC9 varanus pan

17	9	81.8	10	8	Q94VFO	Q94VFO varanus kin
18	9	81.8	10	8	Q9TG47	Q9TG47 ophisaurus
19	9	81.8	10	8	P92771	P92771 xenosaurus
20	9	81.8	10	8	Q94V85	Q94V85 varanus var
21	9	81.8	10	8	Q9TG32	Q9TG32 ophisaurus
22	9	81.8	10	8	Q94PD8	Q94PD8 varanus sca
23	9	81.8	10	8	Q9TG38	Q9TG38 ophisaurus
24	9	81.8	10	8	P92576	P92576 bipes bipor
25	9	81.8	10	8	Q9TG35	Q9TG35 ophisaurus
26	9	81.8	10	8	Q94VD2	Q94VD2 varanus pan
27	9	81.8	10	8	Q9TFV5	Q9TFV5 eublepharus
28	9	81.8	10	8	Q8SHN1	Q8SHN1 bradypodion
29	9	81.8	10	8	Q85J75	Q85J75 varanus bre
30	9	81.8	10	8	Q9C1R7	Q9C1R7 saccharomyc
31	9	81.8	11	7	O77900	O77900 oreochromis
32	9	81.8	11	7	O77917	O77917 oreochromis
33	9	81.8	11	7	O77902	O77902 oreochromis
34	9	81.8	11	7	O77921	O77921 pseudotroph
35	9	81.8	11	7	O77901	O77901 oreochromis
36	9	81.8	11	7	O77916	O77916 oreochromis
37	9	81.8	11	7	O77905	O77905 oreochromis
38	9	81.8	11	7	O77899	O77899 oreochromis
39	9	81.8	11	7	O78121	O78121 oreochromis
40	9	81.8	11	7	O77904	O77904 oreochromis
41	9	81.8	11	7	O77903	O77903 oreochromis
42	9	81.8	11	8	Q94V94	Q94V94 varanus sto
43	9	81.8	11	8	Q9G631	Q9G631 calotes nig
44	9	81.8	11	8	Q9G622	Q9G622 salea horsf
45	9	81.8	11	8	Q94VG8	Q94VG8 varanus gou

ALIGNMENTS

RESULT 1

Q16468 Q16468 PRELIMINARY; PRT; 8 AA.  
AC Q16468;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE DNA for cosmid cC13-1134 PCR primer 1 (Fragment).  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96435920; PubMed=8838806;  
RA James L.A., Ogilvie D.J., Yamakawa K., Nakamura Y., Stirling C.J.,  
RA Anand R.;  
RT "Walking, cloning and mapping with YACs in 3q27. Localisation of 5  
RT ESTs including 3 members of the Cystatin gene family and  
RT identification of CpG islands.";  
RL Genomics 32:425-430(1996).  
DR EMBL; X88976; CAA61407.1; -.  
FT NON TER 8  
SQ SEQUENCE 8 AA; 925 MW; FD5411A7376871E6 CRC64;

Query Match 81.8%; Score 9; DB 4; Length 8;  
Best Local Similarity 66.7%; Pred. No. 1e+06;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3  
Db 5 LTR 7

RESULT 2

Q974Y2 Q974Y2 PRELIMINARY; PRT; 8 AA.  
ID Q974Y2;  
AC Q974Y2;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE COI gene product (Fragment).  
 OS Asterina pectinifera (Starfish).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;  
 OC Asteroidea; Valvatata; Valvatida; Asterinidae; Asterina.  
 OX NCBI\_TaxID=7594;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89354669; PubMed=2766382;  
 RA Jacobs H.T., Asakawa S., Araki T., Miura K., Smith M.J., Watanabe K.;  
 RT "Conserved tRNA gene cluster in starfish mitochondrial DNA.";  
 RL Curr. Genet. 15:193-206(1989).  
 DR ENBL; X16886; CAA34767.1; --  
 DR GO; GO:0005739; C:mitochondrion; IEA.  
 KW Mitochondrion.  
 FT NON\_TER  
 SQ SEQUENCE 8 AA; 1114 MW; F0C9D36415B736D6 CRC64;

Query Match 81.8%; Score 9; DB 8; Length 8;  
 Best Local Similarity 66.7%; Pred. No. 1e+06;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
 Db 3 LSR 5

RESULT 3  
 O89965 PRELIMINARY; PRT; 8 AA.  
 ID O89965;  
 AC O89965;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE Agnoprotein (Fragment).  
 OS Polyomavirus JC.  
 OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.  
 OX NCBI\_TaxID=10632;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98244352; PubMed=9884961;  
 RA Boldorini R., Caldarelli-Stefano R., Monga G., Zocchi M., Mediat M.,  
 RA Tosoni A., Ferrante P.;  
 RT "PCR detection of JC virus DNA in the brain tissue of a 9-year-old  
 child with pleomorphic xanthoastrocytoma.";  
 RL J. Neurovirol. 4:242-245(1998).  
 DR ENBL; AF064547; AAC23995.1; --  
 DR NON\_TER  
 FT NON\_TER  
 SQ SEQUENCE 8 AA; 1002 MW; ED15B736C40732C6 CRC64;

Query Match 81.8%; Score 9; DB 12; Length 8;  
 Best Local Similarity 66.7%; Pred. No. 1e+06;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
 Db 6 LSR 8

RESULT 4  
 Q9T688 PRELIMINARY; PRT; 9 AA.  
 ID Q9T688  
 AC Q9T688;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Cytochrome c oxidase subunit I (Fragment).  
 GN COI  
 OS Gecko gecko (Tokay gecko).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC

OC Lepidosauria; Squamata; Scleroglossa; Gekkota; Gekkonidae; Gekko.  
 OX NCBI\_TaxID=36310;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99343618; PubMed=10413626;  
 RA Macey J.R., Wang Y., Ananjeva N.B., Larson A., Papenfuss T.J.;  
 RT "Vicariant patterns of fragmentation among gekkonid lizards of the  
 RT genus teratocincus produced by the Indian collision: A molecular  
 RT phylogenetic perspective and an area cladogram for central asia.";  
 RL Mol. Phylogenet. Evol. 12:320-332(1999).  
 DR ENBL; AF114249; AAD51600.1; --  
 DR GO; GO:0005739; C:mitochondrion; IEA.  
 KW Mitochondrion.  
 FT NON\_TER  
 SQ SEQUENCE 9 AA; 1188 MW; 428CB9C9D36411A7 CRC64;

Query Match 81.8%; Score 9; DB 8; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1e+06;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
 Db 2 LTR 4

RESULT 5  
 Q94VH4 PRELIMINARY; PRT; 9 AA.  
 ID Q94VH4;  
 AC Q94VH4;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Cytochrome c oxidase subunit I (Fragment).  
 GN COI.  
 OS Varanus glauerti.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.  
 OX NCBI\_TaxID=169841;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ast J.C.;  
 RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";  
 RL Cladistics 17:0-0(2001).  
 DR ENBL; AF407500; AAL10054.1; --  
 DR GO; GO:0005739; C:mitochondrion; IEA.  
 KW Mitochondrion.  
 FT NON\_TER  
 FT NON\_TER  
 SQ SEQUENCE 9 AA; 1124 MW; 9E80C733640DD731 CRC64;

Query Match 81.8%; Score 9; DB 8; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1e+06;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
 Db 3 LAR 5

RESULT 6  
 Q94VD8 PRELIMINARY; PRT; 9 AA.  
 ID Q94VD8;  
 AC Q94VD8;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Cytochrome c oxidase subunit I (Fragment).  
 GN COI  
 OS Varanus niloticus (Nile monitor).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.  
 OX NCBI\_TaxID=62046;

```

RN SEQUENCE FROM N.A.
RP
RA Ast J.C.;
RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL Cladistics 17:0-0(2001).
DR EMBL; AF407514; AAL10096.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON TER
SQ SEQUENCE 9 AA; 1154 MW; 9E80C7336411A731 CRC64;

Query Match      81.8%; Score 9; DB 8; Length 9;
Best Local Similarity 66.7%; Pred. No. 1e+06;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 3 LTR 5

RESULT 7
Q94V18 ID Q94V18 PRELIMINARY; PRT; 9 AA.
AC Q94V18;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome c oxidase subunit I (Fragment).
GN COI.
OS Varanus eremius.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX NCBI_TaxID=62040;
RN [1]
RP SEQUENCE FROM N.A.
RA Ast J.C.;
RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL Cladistics 17:0-0(2001).
DR EMBL; AF407495; AAL10040.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON TER
SQ SEQUENCE 9 AA; 1124 MW; 9E80C733640DD731 CRC64;

Query Match      81.8%; Score 9; DB 8; Length 9;
Best Local Similarity 66.7%; Pred. No. 1e+06;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 3 LAR 5

RESULT 8
Q94VC6 ID Q94VC6 PRELIMINARY; PRT; 9 AA.
AC Q94VC6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome c oxidase subunit I (Fragment).
GN COI.
OS Varanus pilbarensis.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX NCBI_TaxID=62048;
RN [1]
RP SEQUENCE FROM N.A.
RA Ast J.C.;
RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL Cladistics 17:0-0(2001).

```

```

DR EMBL; AF407518; AAL10108.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON TER
SQ SEQUENCE 9 AA; 1064 MW; 874CASA36411A735 CRC64;

Query Match      81.8%; Score 9; DB 8; Length 9;
Best Local Similarity 66.7%; Pred. No. 1e+06;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 3 LTR 5

RESULT 9
Q94VE1 ID Q94VE1 PRELIMINARY; PRT; 9 AA.
AC Q94VE1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome c oxidase subunit I (Fragment).
GN COI.
OS Varanus mertensi (Mertens' water monitor).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX NCBI_TaxID=62044;
RN [1]
RP SEQUENCE FROM N.A.
RA Ast J.C.;
RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL Cladistics 17:0-0(2001).
DR EMBL; AF407512; AAL10090.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON TER
SQ SEQUENCE 9 AA; 1154 MW; 9E80C7336411A731 CRC64;

Query Match      81.8%; Score 9; DB 8; Length 9;
Best Local Similarity 66.7%; Pred. No. 1e+06;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 3 LTR 5

```

```

RESULT 10
Q91BM8 ID Q91BM8 PRELIMINARY; PRT; 9 AA.
AC Q91BM8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Papovavirus BK (Gardner derived clone BKV9) early transcription control region (Fragment).
DE Simian virus 12.
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=46771;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Gardner;
RC MEDLINE=87061221; PubMed=3023684;
RX Chuke W.F., Walker D.L., Peitzman L.B., Frisque R.J.;
RT "Construction and characterization of hybrid polyomavirus genomes.";
RL J. Virol. 60:960-971(1986).
DR EMBL; M14452; AAA96236.1; -.
FT NON TER
SQ SEQUENCE 9 AA; 1130 MW; C7FD15B736C40732 CRC64;

Query Match      81.8%; Score 9; DB 12; Length 9;

```

Best Local Similarity 66.7%; Pred. No. 1e+06;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	LXR	3
Db	6	LSR	8

RESULT 11	Q9PYK1	PRELIMINARY;	PRT;	9 AA.
ID	Q9PYK1			
AC	Q9PYK1			
DT	01-MAY-2000 (TrEMBLrel. 13, Created)			
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)			
DE	Papovavirus BK (Gardner) early transcription control region			
DE	(Fragment).			
OS	Simian virus 12.			
OS	Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.			
OX	NCBI_TaxID:46771;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Gardner;			
RX	MEDLINE=8706121; PubMed=3023684;			
RA	Chuke W.F., Walker D.L., Peitzman L.B., Frisque R.J.;			
RT	"Construction and characterization of hybrid polyomavirus genomes."			
RL	J. Virol. 60:960-971(1986).			
DR	EMBL; M14451; AAA96235.1; -.			
FT	NON TER			
FT	9			
SEQUENCE	9 AA;			
SEQ	C7FD15B736C40732 CRC64;			

RESULT	12
Q958K9	
ID	Q958K9 PRELIMINARY; PRT; 10 AA.
AC	Q958K9;
DT	01-DEC-2001 (TRENBLrel. 19, Created)
DT	01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT	01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE	Cytochrome c oxidase subunit I (fragment).
GN	COI.
OS	Rana boylei.
OG	Mitochondrion.
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Rana.
OX	NBI_TaxID=160499;
OX	(1)
RN	SEQUENCE FROM N.A.
RP	
RX	MEDLINE=21184280; PubMed=11286498;
RA	Macey J.R., Strasburg J.L., Brisson J.A., Vredenburg V.T.,
RA	Jennings M., Larson A.;
RT	"Molecular Phylogenetics of Western North American Frogs of the Rana
RT	boylei Species Group."; RL
RT	Mol. Phylogenet. Evol. 19:131-143(2001).
DR	EMBL; AF314019; AAK56877.1; -
DR	GO; GO:0005739; C.mitochondrion; IEA.
KW	Mitochondrion.
FT	NON_TER 10
SQ	SEQUENCE 10 AA; 1320 MW; 42D380C9D36411A7 CRC64;

D<sub>b</sub> 3 LTR 5

RESULT 13	Q9TG86	PRELIMINARY;	PRT;	10 AA.
ID	Q9TG86			
AC	Q9TG86;			
DT	01-MAY-2000 (TRENBLERel. 13, Created)			
DT	01-MAY-2000 (TRENBLERel. 13, Last sequence update)			
DT	01-JUN-2003 (TRENBLERel. 24, Last annotation update)			
DE	Cytochrome c oxidase subunit I (Fragment).			
GN	COI			
OS	Diploglossus bilobatus.			
OG	Diploglossus.			
OG	Mitochondrion.			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Anguillidae;			
OC	Diploglossus.			
OX	NCBI_TaxID=102183;			
OX	[1]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=93343613; PubMed=10413621;			
RA	Macey J.R., Schulte J.A. II, Larson A., Tuniyev B.S., Orlov N.,			
RA	Papenfuss T.J.;			
RT	"Molecular phylogenetics, tRNA evolution, and historical biogeography			
RT	in anguillid lizards and related taxonomic families.";			
RL	Mol. Phylogenet. Evol. 13:250-272 (1999).			
DR	EMBL; AF085608; AAD51514.1; "			
DR	GO; GO:0005739; C:mitochondrion; IEA.			
KW	Mitochondrion.			
FT	NON_TER			
FT	10			
FT	SEQUENCE	10 AA;	1235 MW;	5DEE80C7336411A7 CRC64;
SQ				
Query Match		81.8%;	Score 9;	DB 8; Length 10;
Best Local Similarity		56.7%;	Pred. No. 4.7e+03;	
Matches 2;	Conservative	0;	Mismatches 1;	Indels 0;
				Gaps 0;

RESULT 14	
Q94V97	PRELIMINARY; PRT; 10 AA.
Q94V97	
AC	
Q1-DEC-2001	(T-EMBLrel. 19, Created)
DT	
Q1-DEC-2001	(T-EMBLrel. 19, Last sequence update)
DT	
Q1-JUN-2003	(T-EMBLrel. 24, Last annotation update)
DT	
DE	Cytochrome c oxidase subunit I (Fragment).
DE	
COI	
OS	Varanus spenceri.
OS	
OC	Mitochondrion.
OC	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX	NCB1_TaxID=169854;
OX	
[1]	
SEQUENCE FROM N.A.	
AST J.C.;	
ERA	"Mitochondrial DNA evidence and evolution in Varanicoidea (Squamata).";
ERA	
RL	Cladistics 17:0-0(2001).
RL	
EMBL	AF407530; AAL10142.1; -.
DR	GO; GO:0005739; C.mitochondrion; IEA.
DR	
KW	Mitochondrion.
NON TER	10
SEQUENCE	10 AA; 1255 MW; SDEE80C7336411A7 CRC64;
SEQ	

Query Match	Score 9;	DB 8;	Length 10;	Score 9;	DB 8;	Length 10;
Best Local Similarity	81.8%;			81.8%;		
Pred. No. 4.7e+03;				Pred. No. 4.7e+03;		
Matches 2;				Matches 2;		
Conservative				Conservative		
Mismatches 0;				Mismatches 0;		
Indels 0;				Indels 0;		
Gaps 0;				Gaps 0;		

Db 3 LTR 5

RESULT 15

Q94VDS PRELIMINARY; PRT; 10 AA.  
AC Q94VDS;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
DE Cytochrome c oxidase subunit I (Fragment).  
GN COI.  
OS Varanus olivaceus (Gray's monitor).  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.  
OX NCBI\_TaxID=62047;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Ast J.C.;  
RL "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";  
RL Cladistics 17:0-0(2001).  
DR EMBL; AF407515; AAL10099.1; -.  
DR GO; GO:0005739; C:mitochondrion; IEA.  
KW Mitochondrion.  
FT NON TER 10 10  
SQ SEQUENCE 10 AA; 1234 MW; 584C4CA5A36411A7 CRC64;

Query Match 81.8%; Score 9; DB 8; Length 10;  
Best Local Similarity 66.7%; Pred. No. 4.7e+03;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3  
|  
Db 3 LTR 5

Search completed: September 24, 2004, 07:30:28  
Job time : 7.16842 secs

This page Blank (usp10)